ملخصات شوم

نظريات ومسائل

في

الإحصاء والاقتصاد القياسى


الله
بسم الله الرحمن الرحيم

ملخصات شوم

نظريات ومسائل

في

الإحصاء والاقتصاد القياسي

تأليف

دومينيك سالفاتور Ph.D

أستاذ الاقتصاد

جامعة فوردهام

ترجمة

الدكتورة سعدية حافظ منتصر

قسم الإحصاء - كلية التجارة

جامعة عين شمس - جمهورية مصر العربية

مراجعة

الأستاذ الدكتور عبد العظيم أنيس

أستاذ غير متفرغ - جامعة عين شمس

جمهورية مصر العربية



الدار الدولية للنشر والتوزيع

بسم الله الرحمن الرحيم

مقدمة الناشر

المعرفة هى أصل الحضارة ،
والكلمة هى مصدر المعرفة ،
والكلمة المطبوعة هى أهم مكون فى هذا المصدر .

وقد كانت الكلمة المطبوعة ولاتزال أهم وسائل الثقافة والاعلام وأوسعها انتشاراً وأبقاها اثراً ، حيث حملت إلينا حضارات الأمم عبر آلاف السنين لتتولى الأجيال المتلاحقة صياغة حضاراتها وإضاءة الطريق بنور العلم والمعرفة

والكلمة تبقى مجرد فكرة لدى صاحبها حتى تتاح لها فرصة نشرها وترجمتها إلى لغات الآخرين ثم توزيعها ، وذلك وحده هو الذى يكفل لها أداء رسالتها .

وعالم الكتب العلمية عالم رحب ممتد الآفاق ، متسع الجنبات ، والعلم لا وطن له ولا حدود . ويوم يحظى القارىء بأحدث الكتب العلمية باللغة العربية لهو اليوم الذى تتطلع له الأمة العربية جمعاء .

والدار الدولية للنشر والتوزيع تشعر بالرضا عن مساهمتها فى هذا المجال بتقديم الطبعات العربية للكتب العلمية الصادرة عن دار ماكجروهيل للنشر بموجب الاتفاق المبرم معها ، مستهدفة توفير احتياجات القارىء العربى أستاذاً وباحثاً وممارساً .

ومن جانب آخر فنحن نمد يدنا إلى الجامعات العربية والمراكز العلمية والمؤسسات والهيئات الثقافية للتعاون معنا فى إصدار طبعات عربية حديثة من الكتب والمراجع العلمية تخدم التقدم العلمى والحضارى للقارىء العربى

والله ولى التوفيق

محمد وفائى كامل
مدير عام
الدار الدولية للنشر والتوزيع

مقدمة

يقدم هذا الكتاب مدخلا واضحاً ومحكماً فى الإحصاء والاقتصاد القياسى . وغالباً ماتكون مادة الإحصاء أو الاقتصاد القياسى واحدة من أكثر المواد صعوبة وأكثرها فائدة فى نفس الوقت ، من بين المواد التى تدرس بالكليات والجامعات . ومن ثم ، فإن الهدف من هذا الكتاب هو التغلب على هذه الصعوبة باستخدام مدخل يعتمد على الشرح من خلال حل المسائل .

ويبدأ كل فصل بتقرير النظرية ، والمبادىء أو الخلفية اللازمة من المعلومات ، موضحة تماماً بأمثلة . ثم يتبع ذلك العديد من المسائل النظرية والعملية مصحوبة بحلول تفصيلية « خطوة بعد خطوة » . وبينما قصد أساساً من هذا الكتاب أن يكون مكملا للكتب القياسية المتداولة فى الإحصاء والاقتصاد القياسى ، فإنه يمكن استخدامه أيضاً كمرجع مستقل بذاته أو بالإضافة إلى المحاضرات .

ويقدم الكتاب مادة فى الإحصاء والاقتصاد القياسى تكفى فصلا دراسياً واحداً أوسنة كاملة لطلاب الجامعة فى الاقتصاد ، إدارة الأعمال ، والعلوم الاجتماعية . كما أنه يقدم مرجعاً مفيداً جداً لطلاب الماجستير ولكل من يستخدمون أويرغبون فى استخدام الإحصاء والاقتصاد القياسى فى أعمالهم . وهو لايفترض خلفية إحصائية لدى القارىء .

ويعتبر الكتاب متكاملا من حيث أنه يغطى مواد الإحصاء (الفصول ١ — ٥) المطلوبة لدراسة الاقتصاد القياسى (الفصول ٦—١٠) ، ويركز الكتاب على الجانب التطبيقى حيث تأتى البراهين فى قسم المسائل وليس فى سياق الشرح . كما يستخدم الكتاب كلما أمكن بيانات واقعية اقتصادية — اجتماعية وفى مجال الأعمال لتوضيح أساليب ونماذج الاقتصاد القياسى الأكثر تقدماً . وقد تضمن الكتاب برنامج كمبيوتر كامل لتوضيح كيفية استخدام وتفسير النتائج باستخدام واحد من أكثر البرامج الإحصائية شيوعاً :
The Statistical Package for the Social Sciences (*SPSS*) وقد تعرض الكتاب بوضوح وأحكام لموضوعات فى الاقتصاد القياسى تعترض الباحثين كثيراً ، مثل تعدد العلاقات الخطية والارتباط الذاتى حيث يناقش المشاكل الناجمة عنها وطرق اختبار وجودها والأساليب الممكنة لتصويبها . كما يتضمن الكتاب عينة من امتحانات الاحصاء والاقتصاد القياسى .

وقد تم اختبار منهجية هذا الكتاب والكثير من محتوياته عند تدريس الإحصاء والاقتصاد القياسى على مستوى البكالوريوس والدراسات العليا بجامعة فورد هام ، ووجد الطلاب أن منهجية ومحتويات الكتاب بالغة الفائدة وقدموا اقتراحات قيمة عديدة لتحسينه . كما تلقيت اقتراحات مفيدة جداً من الأساتذة جون بيدريه وإدوارد داولنج من جامعة فورد هام . كما أن الطلاب التالية أسماؤهم قد قرأوا بعناية أصول الكتاب وقدموا الكثير من الاقتراحات المفيدة : وليم فوت ، فرانك التيرى ، سيسليا ونترز ، توماس لودر ، ريتشارد مايكلفيلدر أنيتا باسمانتير وكونى وامورين رايز . وإليهم جميعاً أقدم عميق امتنانى . كما أننى مدين علمياً لأساتذتى السابقين فى الإحصاء والاقتصاد القياسى : جاك جونستون ، لورانس كلاين وبرنارد أوكن . وأخيراً ، أود أن أعبر عن امتنانى لجوزيف ميدلتون ، وتشارلز بارسيلونا ، ومارى جرير من مركز كمبيوتر جامعة فورد هام وكذلك للعاملين بسلسلة شوم فى دار ماكجروهيل للنشر ، خصوصاً جون آليانو ونك مونتى لمساعداتهم الطيبة والماهرة .

كما أننى ممتن لمنفذ وصية المرحوم سير رونالد أ . فيشر وللدكتور فرانك ياتس ومجموعة لونجمان لميتد ، لندن ، لسماحهم باستخدام وتعديل الجداول ٣ و ٤ من كتابهم « جداول إحصائية للبحوث البيولوجية والزراعية الطبية » .

إن سلسلة شوم فى الاقتصاد تتضمن بالإضافة إلى كتاب الإحصاء والاقتصاد القياسى الكتب التالية : **نظرية التحليل الجزئى ، نظرية التحليل الكلى ، الاقتصاد الدولى ، اقتصاديات التنمية ، الرياضيات للاقتصاديين ، وأصول الاقتصاد .**

دومينيك سالفاتور

نيويورك ١٩٨٢

المحتويات

الفصل الأول

تمهيـــد

١ـ١ طبيعة علم الاحصاء

يختص علم الإحصاء بجمع وعرض وتحليل واستخدام البيانات الرقية لعمل استدلالات واتخاذ قرارات فى ظل عدم التأكد فى مجالات الاقتصاد والأعمال وغيرها من العلوم الاجتماعية والطبيعية .

وينقسم الإحصاء إلى الإحصاء الوصفى والإحصاء الاستدلالى . ويختص **الإحصاء الوصفى** بتلخيص وتوصيف مجموعة من البيانات . بينما يختص **الإحصاء الاستدلالى** بالوصول إلى تعميم عن خواص الكل (ويسمى **المجتمع**) من واقع فحص جزء من هذا الكل (ويسمى **العينة**) ولـكى يكون هذا التعميم سليما فإن العينة يجب أن تكون ممثلة للمجتمع وأن يتم تحديد احتمال الخطأ فى هذا التعميم .

وسوف نناقش الإحصاء الوصفى بالتفصيل فى الفصل الثانى ، ويلى ذلك الاستدلال الإحصائى (وهو الأهم) ، حيث نتناول موضوع الاحتمال فى الفصل الثالث وموضوع التقدير فى الفصل الرابع واختبارات الفروض فى الفصل الخامس .

مثال (١) : افترض أن لدينا بيانات عن دخل 1000 أسرة أمريكية . فإن هذه البيانات يمكن تلخيصها بإيجاد متوسط دخل الأسرة وتعيين مدى انتشار دخل الأسر حول هذا المتوسط . ويمكن أيضاً توصيف البيانات بإنشاء جدول أو رسم بيانى لعدد أو نسبة العائلات فى كل فئة من فئات الدخل . أن هذا مانعنيه بالإحصاء الوصفى . أما إذا كانت هذه الأسر (1000 أسرة) ممثلة لجميع العائلات الأمريكية فإنه يمكننا تقدير متوسط لدخل الأسرة فى الولايات المتحدة كلها وإجراء اختبارات للفروض عن هذا المتوسط . وحيث أن هذه النتائج معرضة للخطأ فإن علينا أيضاً أن نحدد احتمال الخطأ فى هذه النتائج . إن هذا هو موضوع الاستدلال الإحصائى .

١ـ٢ الاحصاء والاقتصاد القياسى

يختص **الاقتصاد القياسى** بتطبيق النظرية الاقتصادية ، والرياضيات ، والأساليب الإحصائية فى اختبار الفروض ، والتقدير ، والتنبؤ بالظواهر الاقتصادية . وقد ارتبط الاقتصاد القياسى ارتباطاً وثيقاً **بتحليل الانحدار** . وينصب تحليل الانحدار على قياس العلاقة بين متغير تابع ومتغير مستقل أو أكثر . وحيث أن العلاقات بين المتغيرات الاقتصادية تكون بصفة عامة غير دقيقة فإنه يجب إضافة عنصر يمثل الخطأ أو التشويش (له خواص احتمالية محددة) فى العلاقة (أنظر المسألة ١ ـ ٨) .

ويتعلق الفصلان السادس والسابع بتحليل الانحدار ، ويمثل الفصل الثامن امتداد نموذج الانحدار الأساسى ، ويتناول الفصل التاسع طرقاً لاختبار فروض نموذج الانحدار الأساسى ولاجراء التصحيحات الناتجة عن الخروج على فروض النموذج فى حين يتناول الفصل العاشر طرق التقدير للمعادلات الآنية . ومن ثم فإن الفصول من (١ ـ ٥) تتناول الإحصاء اللازم لدراسة الاقتصاد القياسى (الفصول من ٦ ـ ١٠) .

مثال (٢) : تخبرنا نظرية الاستهلاك أن الناس عموماً يزيدون من إنفاقهم على الاستهلاك C كلما زاد الدخل (بعد الضرائب) المتاح Y_d ، ولكن الزيادة فى الاستهلاك لاتكون بنفس قدر الزيادة فى الدخل المتاح . ويمكن التعبير عن ذلك بمعادلة خطية صريحة كالآتى :

$$C = b_0 + b_1 Y_d \qquad (١ - ١)$$

حيث b_0 و b_1 ثوابت مجهولة تسمى معالم . فالمعلمة b_1 هي ميل خط الانحدار وتمثل الميل الحدى للاستهلاك MPC . وحيث أنه من المرجح حتى بالنسبة للأفراد الذين تتساوى دخولهم المتاحة أن يختلف إنفاقهم الاستهلاكى ، فإن العلاقة الدقيقة نظرياً والمحددة بالمعادلة (١ - ١) يجب أن تعدل بإضافة عنصر تشويش عشوائى أو حد الخطأ u بحيث تكون المعادلة ذات طابع احتمالى على النحو التالى .

$$C = b_0 + b_1 Y_d + u \qquad (٢ - ١)$$

١-٢ منهاج الاقتصاد القياسى

تتضمن بحوث الاقتصاد القياسى ، بصفة عامة ، المراحل الثلاث الآتية :

المرحلة ١ : تحديد النموذج أو الفرض المستخدم فى شكل معادلة احتمالية صريحة ، مع توقعات نظرية مسبقة عن إشارة وحجم معالم الدالة .

المرحلة ٢ : جمع بيانات عن متغيرات النموذج وتقدير معاملات الدالة باستخدام أساليب الاقتصاد القياسى المناسبة (الفصول من ٦ إلى ٨) .

المرحلة ٣ : تقويم المعاملات المقدرة فى الدالة باستخدام معايير الاقتصاد والإحصاء والاقتصاد القياسى .

مثال (٣) : المرحلة الأولى لبحوث الاقتصاد القياسى فى نظرية الاستهلاك تكون بتقديم النظرية فى شكل معادلة احتمالية صريحة ، كما فى معادلة (١ - ١) ، مع توقع أن تكون $b_0 > 0$ (أى أنه عند $Y_d = 0$ فإن $C > 0$ إذ أن المستهلك يسحب من مدخراته أو يقترض لكى يستهلك) ، وأن $0 < b_1 < 1$. وتتضمن المرحلة الثانية جمع بيانات عن الإنفاق الاستهلاكى والدخل المتاح واستخدامها فى تقدير المعادلة (١ - ١) . وتتضمن المرحلة الثالثة فى بحوث الاقتصاد القياسى (١) التأكد عما إذا كانت القيمة المقدرة $b_0 > 0$ ، والقيمة المقدرة $0 < b_1 < 1$ ، (٢) تحديد ما إذا كانت نسبة « مرضية » من التغير فى C يمكن تفسيرها « كنتيجة للتغير فى Y_d » ، وكذلك ما إذا كانت كل من b_0 ، b_1 « معنوية إحصائياً عند مستوى معنوية مقبول » (أنظر المسألة ١ - ١٣ (حـ) والقسم ٥ - ٣) ، (٣) اختبار ما إذا كانت شروط نموذج الانحدار الأساسى متوافرة ، فإن لم تتوافر ، يحدد كيفية إجراء تصحيح نتيجة الخروج على هذه الشروط . فإذا لم تجتز العلاقة المقدرة هذه الاختبارات ، فيجب تعديل العلاقة المفترضة وإعادة التقدير حتى يتم التوصل إلى علاقة استهلاك مقدرة مرضية

مسائل محـــلولة

طبيعة عـــلم الإحصاء :

١ - ١ ماهو الفرض وما هى وظيفة كل من (أ) مجال دراسة الإحصاء ؟ (ب) الإحصاء الوصفى ؟ (حـ) الاستدلال الإحصائى ؟

(أ) الإحصاء مجموعة من الإجراءات والأساليب المستخدمة فى جمع وعرض تحليل البيانات التى تبنى عليها القرارات فى مواجهة عدم التأكد أو فى مواجهة معلومات ناقصة . وفى الوقت الحاضر نجد أن التحليل الإحصائى يستخدم تقريباً فى كل مهنة . فالاقتصادى يستخدمه لاختبار كفاءة أساليب الإنتاج المختلفة ، ورجل الأعمال قد يستخدمه لاختبار تصميم أو تغليف المنتج بما يعظم المبيعات ، والباحث الاجتماعى يستخدمه لتحليل نتائج عقار معين على برنامج تأهيل ، وعالم النفس الصناعى لدراسة استجابات العمال لظروف العمل بالمصنع ، والعالم السياسى للتنبؤ بأنماط التصويت ، والطبيب لاختبار فعالية عقار جديد ، والكيميائى لإنتاج أسمدة أرخص ، وهكذا . .

(ب) الإحصاء الوصفى يختزل مجموعة البيانات إلى معلومة أو اثنتين تميزان كل البيانات . ويتعلق أيضاً الإحصاء الوصفى بتقديم مجموعة البيانات على شكل جداول أو رسوم بيانية ، وغيرها من وسائل العرض البيانى .

(حـ) الاحصاء الاستدلالى (ويشمل التقدير واختبارات الفروض) ويتعلق باستخلاص تعميمات عن خواص المجتمع من واقع خواص عينة مأخوذة من هذا المجتمع . ومن ثم فإن الإحصاء الاستدلالى يتضمن تعليلا استقرائياً . (وذلك على نقيض التعليل الاستنباطى الذى يستنبط خواص الجزء مبتدئاً بالكل) .

١ – ٢ (أ) أيها أكثر أهمية فى الوقت الحاضر ؟ الإحصاء الوصفى أم الإحصاء الاستدلالى ؟
(ب) ماهى أهمية استخدام عينة ممثلة فى الاستدلال الإحصائى ؟
(ﺤ) لماذا نحتاج نظرية الاحتمالات ؟

(أ) بدأ الإحصاء كعلم وصفى بحت ، ولكنه تطور إلى أداة قوية لاتخاذ القرارات مع نمو فرع الاستدلال فيه . وأصبح التحليل الإحصائى الحديث ينصب أساساً على الإحصاء الاستدلالى . ومع ذلك فإن الإحصاء الاستدلالى والإحصاء الاستنباطى مكملان أحدهما للآخر . وقبل أن نتعلم التعميم من العينات إلى المجتمعات يجب أن نتعلم كيفية توليد العينات من المجتمع .

(ب) لكى يكون الاستدلال الإحصائى سليما يجب أن يستند إلى عينة تعكس تماماً صفات وخواص المجتمع الذى سحبت منه . وتكون العينة ممثلة إذا كانت المعاينة عشوائية حيث يكون لكل مفردة من مفردات المجتمع نفس الفرصة للدخول فى العينة (أنظر قسم ٤ – ١) .

(ﺤ) حيث أن احتمال الخطأ وارد فى الاستدلال الإحصائى فإن تقديرات واختبارات خواص المجتمع تعطى ومعها فرصة أو احتمال الخطأ فى هذه التقديرات أو الاختبارات ومن هنا فإن نظرية الاحتمالات تعتبر عنصراً أساسياً فى الاستدلال الإحصائى .

١ – ٣ كيف يمكن لمدير شركة تنتج مصابيح كهربائية أن يلخص ويصف لاجتماع مجلس الإدارة نتائج اختبار عمر عينة من 100 مصباح من إنتاج الشركة ؟

ان عرض البيانات (الخام) عن عمر كل مصباح فى العينة أمر غير ملائم ويستغرق وقتاً طويلا من أعضاء المجلس لتقويمها . ويمكن بدلا من ذلك أن يقوم المدير باختزال البيانات بتوضيح أن متوسط عمر المصابيح التى تم اختبارها فى العينة هو 360 ساعة وأن 95% من المصابيح التى فحصت قد عاشت بين 320 و 400 ساعة . ويكون المدير بذلك قد قدم معلومتين (متوسط العمر وانتشار المفردات حول القيمة الوسطى) تميزان عمر المصابيح المائة التى تم فحصها . وقد يرى المدير أيضاً أن يصف البيانات باستخدام جدول أو رسم بيانى يوضح عدد أو نسبة المصابيح موزعة على فئات طول كل منها عشرة ساعات طبقاً للعمر الذى قضاه كل مصباح . ويكون مثل هذا العرض الجدولى أو البيانى مفيداً فى الإلمام العام السريع بالبيانات . وبتلخيص وتوصيف البيانات على النحو الموضح يكون المدير قد استخدم الإحصاء الوصفى . وجدير بالذكر أن الإحصاء الوصفى يمكن استخدامه فى تلخيص وتوصيف أى مجموعة من البيانات سواء كانت عينة (كالمثال السابق) أو مجتمعاً (عندما تكون مفردات المجتمع معروفة ويمكن قياس خواصها) .

١ – ٤ (أ) لماذا قد يرغب المدير فى مسألة ١ – ٣ أن يتطرق إلى الاستدلال الإحصائى ؟
(ب) ماذا يتضمن هذا وماذا يتطلب ؟

(أ) تتطلب مراقبة جودة الإنتاج أن يكون لدى المدير فكرة جيدة تماماً عن متوسط عمر المصابيح الكهربائية التى تنتجها الشركة والانتشار حول هذا المتوسط . غير أن فحص جميع المصابيح الكهربائية يؤدى إلى تدمير إنتاج الشركة كله . وحتى عندما لايؤدى الفحص إلى تدمير المنتج ، فإن فحص الإنتاج كله يكون عادة باهظ التكلفة ويستغرق وقتاً طويلا . وعليه ، فإن الإجراء المتبع هو أخذ عينة من الإنتاج والاستدلال على خواص وصفات الإنتاج كله (المجتمع) من الصفات المناظرة للعينة المسحوبة من المجتمع .

(ب) يتطلب الاستدلال الإحصائى أولا أن تكون العينة ممثلة للمجتمع الذى تؤخذ منه . فإذ كانت الشركة تنتج المصابيح الكهربائية فى مصانع مختلفة ، باستخدام أكثر من وردية واحدة ، وباستخدام مواد خام مشتراة من أكثر من مورد فهذه جميعاً يجب أن تمثل فى العينة بنسبة مساهمتها فى الإنتاج الكلى للشركة . فباستخدام متوسط عمر المصابيح فى العينة والانتشار حول هذا المتوسط يمكن لمدير الشركة أن يقدر ، باحتمال 95% أن يكون تقديره صحيحاً واحتمال 5% أن يكون

تقديره خاطئاً ، أن متوسط العمر لكل المصابيح التى تنتجها الشركة يقع بين 320 و 400 ساعة (أنظر قسم ٤ – ٣) . وكبديل يمكن للمدير أن يستخدم معلومات العينة لكى يختبر ، باحتمال 95% أن يكون على صواب ، واحتمال 5% أن يكون على خطأ ، أن متوسط العمر فى مجتمع جميع المصابيح التى تنتجها الشركة أكبر من 320 ساعة (أنظر قسم ٥ – ٢) . وسواء فى التقدير أو فى اختبار متوسط المجتمع باستخدام بيانات العينة يكون المدير مستخدماً للاستدلال الإحصائى .

الإحصاء والاقتصاد القياسى :

١ – ٥ ماذا يقصد بالآتى (أ) اقتصاد قياسى ؟ (ب) تحليل الانحدار ؟ (ح) حد التشويش أو الخطأ ؟ (د) نماذج المعادلات الآنية ؟

(أ) **الاقتصاد القياسى** هو تكامل للنظرية الاقتصادية مع الرياضيات والأساليب الإحصائية بهدف اختبار فروض عن الظواهر الاقتصادية ، وتقدير معاملات العلاقات الاقتصادية أو التنبؤ بالقيم المستقبلة للمتغيرات أو الظواهر الاقتصادية . ويقسم الاقتصاد القياسى إلى جزءين : نظرى وتطبيق . وبصفة عامة يتضمن **الاقتصاد القياسى النظرى** طرق قياس العلاقات الاقتصادية ، بينما يبحث **الاقتصاد القياسى التطبيقى** المشاكل والنتائج فى مجال اقتصاد معين ، مثل نظرية الطلب الإنتاج ، الاستثمار ، الاستهلاك ، وغيرها من مجالات بحوث الاقتصاد التطبيقية . وعلى أية حال ، فإن الاقتصاد القياسى هو من ناحية علم ومن ناحية أخرى فن إذ أن الحدس والحكم الجيد للباحث يلعبان غالباً دوراً حاسماً .

(ب) يبحث تحليل الانحدار العلاقة السببية بين متغير اقتصادى يحتاج إلى تفسير (المتغير التابع) ومتغير آخر أو أكثر من المتغيرات المستقلة أو المتغيرات التفسيرية . وعندما يكون هناك متغير مستقل أو تفسيرى واحد نكون فى **مجال الانحدار البسيط** . بينما فى الحالات الأكثر شيوعاً عندما يستخدم أكثر من متغير مستقل أو تفسيرى فإننا نكون فى **مجال الانحدار المتعدد** .

(ح) أن التشويش (العشوائى) أو الخطأ يجب أن تتضمنه العلاقات الدقيقة التى تفترضها النظرية الاقتصادية والاقتصاد الرياضى حتى تكون هذه العلاقات ذات طابع احتمالى . (ليعكس ذلك حقيقة أن العلاقات الاقتصادية بين المتغيرات الاقتصادية هى فى العالم الحقيقى غير دقيقة وإلى حد ما شاذة) .

(د) نماذج المعادلات الآنية تشير إلى العلاقات بين المتغيرات الاقتصادية معبراً عنها بأكثر من معادلة وبحيث تتبادل المتغيرات الاقتصادية التأثير فى المعادلات المختلفة . وتعتبر نماذج المعادلات الآنية أكثر نماذج الاقتصاد القياسى تعقيداً ويتم تناولها فى الفصل العاشر .

١ – ٦ (أ) ماهى وظائف الاقتصاد القياسى ؟ (ب) ماهى نواحى الاقتصاد القياسى (وغيره من العلوم الاجتماعية) التى تجعله يختلف اختلافاً أساسياً عن معظم العلوم الطبيعية ؟

(أ) للاقتصاد القياسى ثلاث وظائف متداخلة ، الأولى هى اختبار النظريات أو الفروض الاقتصادية . على سبيل المثال ، هل الاستهلاك مرتبط مباشرة بالدخل ؟ هل هناك علاقة عكسية بين الكمية المطلوبة من سلعة ما وبين سعرها ؟ والوظيفة الثانية للاقتصاد القياسى هى أنه يمد الباحث بتقديرات رقمية لمعاملات العلاقات الاقتصادية . وهذه التقديرات أساسية عند اتخاذ القرار . فعلى سبيل المثال ، فإن متخذ القرار فى الحكومة يحتاج إلى تقدير دقيق لمعاملات العلاقة بين الاستهلاك والدخل لكى يحدد التأثير المتوقع (المضاعف) لخفض مقترح للضرائب . والمدير يحتاج أن يعرف ما إذا كان تخفيض السعر يؤدى إلى زيادة أو نقصان الإيرادات الكلية للمبيعات بالشركة وبأى قدر . والوظيفة الثالثة للاقتصاد القياسى هى التنبؤ بالأحداث الاقتصادية . وهذا أيضاً ضرورى لمتخذ القرار لكى يتخذ الإجراءات التصحيحية المناسبة إذا كان من المتوقع أن يرتفع معدل البطالة أو التضخم فى المستقبل .

(ب) هناك اختلافان أساسيان بين الاقتصاد القياسى (وغيره من العلوم الاجتماعية) من ناحية ومعظم العلوم الطبيعية (مثل الفيزياء) من ناحية أخرى . الاختلاف الأول (كما ذكر من قبل) هو أن العلاقات بين المتغيرات الاقتصادية هى

علاقات غير دقيقة وإلى حد ما شاذة . والاختلاف الثانى أن معظم الظواهر الاقتصادية تحدث متزامنة ومن هنا لايمكن إجراء تجارب معملية عليها . وهذه الاختلافات تتطلب طرقاً خاصة في التحليل (مثل إضافة حد التشويش أو الخطأ إلى العلاقات الدقيقة التى تفترضها النظرية الاقتصاديه) وفى التحليل المتعدد المتغيرات (مثل تحليل الانحدار المتعدد) . وهذا الأخير يعزل تأثير كل متغير مستقل أو تفسيرى على المتغير التابع عند مواجهة تغير متزامن فى كل المتغيرات التفسيرية .

١ – ٧ بأى شكل ولأى غرض تجتمع المجالات الآتية لتكون معاً مجال الاقتصاد القياسى

(أ) النظرية الاقتصادية (ب) الرياضيات (ﺣ) التحليل الإحصائى ؟

(أ) يفترض الاقتصاد القياسى مسبقاً وجود مجموعة من النظريات أو الفروض الاقتصادية التى تحتاج إلى اختبار . فإذا كانت المتغيرات التى تقترحها النظرية الاقتصادية لاتعطى تفسيراً مرضياً فيمكن للباحث تجربة صياغات ومتغيرات بديلة قد تكون ناجمة عن اختبارات سابقة أو نظريات معارضة . ومن هنا فإن بحوث الاقتصاد القياسى يمكن أن تؤدى إلى قبول أو رفض أو إعادة صياغة النظريات الاقتصادية .

(ب) تستخدم الرياضيات للتعبير عن التقريرات اللفظية للنظريات الاقتصادية فى صورة رياضية ، وذلك بالتعبير – على شكل علاقات دالية دقيقة أو محددة – عن العلاقات بين المتغير التابع وبين واحد أو أكثر من المتغيرات المستقلة أو التفسيرية .

(ﺣ) يستخدم التحليل الإحصائى الأساليب الملائمة لتقدير علاقات غير دقيقة وغير قابلة للتجريب بين المتغيرات الاقتصادية باستخدام البيانات الاقتصادية المناسبة ولتقويم النتائج .

١ – ٨ بماذا تبرر إضافة حد التشويش أو الخطأ إلى تحليل الانحدار ؟

إن إضافة حد التشويش (العشوائى) أو حد الخطأ (بمواصفات احتمالية معرفة بدقة) مطلوب فى تحليل الانحدار لثلاثة أسباب هامة . الأول ، طالما أن الغرض من النظرية هو التعميم والتبسيط ، فإن العلاقات الاقتصادية عادة تتضمن فقط أهم القوى المؤثرة . ويعنى هذا أن العديد من المتغيرات الأخرى ذات التأثير الضعيف أو غير المنتظم لاتدخل فى الحسبان . فيمكن النظر إلى حد الخطأ على أنه يمثل التأثير الصافى لعدد كبير من القوى ذات التأثير الصغير أو غير المنتظم . والسبب الثانى ، أنه يمكن تبرير إضافة حد الخطأ بأنه يأخذ فى الاعتبار التأثير الصافى للأخطاء الممكنة فى قياس المتغير التابع أو المتغير الذى يتم تفسيره . وأخيراً حيث أن السلوك الإنسانى ، فى ظل ظروف متطابقة ، يتباين عادة بصورة عشوائية فإن حد التشويش أو الخطأ يمكن استخدامه للإمساك بهذا السلوك . أى أن عنصر الخطأ هذا يسمح بانحرافات فردية عشوائية عن العلاقات المحددة الدقيقة التى تقترحها النظرية الاقتصادية والاقتصاد الرياضى .

١ – ٩ تنص نظرية طلب المستهلك على أن الكمية المطلوبة من سلعة ما D_X ، هى دالة فى سعرها P_X ، ودخل المستهلك Y ، وأسعار سلع أخرى (بديلة أو مكملة) مثلا السعر P_Z للسلعة Z . بافتراض أن أذواق المستهلكين ثابتة خلال فترة الدراسة فالمطلوب التعبير عن النظرية السابقة (أ) فى صورة خطية أو معادلة صريحة أو محددة (ب) فى صورة عشوائية (ﺣ) ماهى المعاملات التى يجب تقديرها وماذا تسمى ؟

(أ) (١ – ٣) $$D_X = b_0 + b_1P_X + b_2Y + b_3P_Z$$

(ب) (١ – ٤) $$D_X = b_0 + b_1P_X + b_2Y + b_3P_Z + u$$

(ﺣ) المعاملات التى يلزم تقديرها هى b_0 . b_1 ، b_2 و b_3 . وتسمى بالمعالم .

منهاج الاقتصـــاد القيـــاسى :

١ – ١٠ بالإشارة إلى نظرية طلب المسهلك فى مسألة ١ – ٩ وضح (أ) ماهى الخطوة الأولى فى بحث الاقتصاد القياسى (ب) ماهى التوقعات النظرية المسبقة لإشارة وحجم المعالم فى دالة الطلب المذكورة فى المعادلة (١ – ٤) .

(أ) الخطوة الأولى فى التحليل الاقتصادى القياسى هى التعبير عن نظرية طلب المستهلك فى صورة معادلة احتمالية كما فى معادلة (١ - ٤) ، ثم تحديد التوقعات النظرية المسبقة عن إشارة و حجم معالم الدالة .

(ب) تفترض نظرية طلب المستهلك أنه فى المعادلة (١ - ٤) ، $b_1 < 0$ (بمعنى أن العلاقة بين السعر و الكمية هى علاقة عكسية) ، وأن $b_2 > 0$ إذا كانت السلعة عادية (بمعنى أن المستهلك يشترى أكثر من السلعة عند مستوى الدخل الأعلى) ، $b_3 > 0$ إذا كانت X و Z سلعاً بديلة ، و $b_3 < 0$ إذا كانت X و Z سلعاً مكملة .

١ - ١١ اذكر المرحلة الثانية لبحث الاقتصاد القياسى (أ) بصفة عامة (ب) فيما يتعلق بدالة الطلب المحددة فى معادلة (١ - ٤) .

(أ) المرحلة الثانية فى بحث الاقتصاد القياسى تتضمن جمع البيانات عن المتغير التابع وعن كل من المتغيرات المستقلة أو التفسيرية فى النموذج واستخدام هذه البيانات للتقدير العملى لمعالم النموذج . ويتم هذا عادة باستخدام تحليل الانحدار المتعدد (ونتناوله فى الفصل السابع) .

(ب) لتقدير دالة الطلب فى المعادلة (١ - ٤) ، يجب تجميع بيانات عن (١) كمية طلب المستهلكين على السلعة X ، (٢) سعر السلعة X ، (٣) دخول المستهلكين ، و (٤) سعر السلعة Z فى وحدة الزمن (أى فى اليوم ، أو الشهر ، أو السنة) وعلى مدى الأيام أو الشهور أو السنوات . ويتم تقدير علاقة انحدار D_x على P_x ، Y و P_z من خلال تقدير المعالم b_0 ، b_1 ، b_2 و b_3 .

١ - ١٢ كيف يمكن أن تختلف نوعية البيانات المطلوبة لتقدير دالة الطلب المذكورة فى معادلة (١ - ٤) عن نوعية البيانات اللازمة لتقدير دالة الاستهلاك لمجموعة من العائلات عند نقطة زمنية معينة ؟

لتقدير معادلة الطلب فى المعادلة (١ - ٤) ، يلزمنا بيانات عن قيم المتغيرات خلال فترة زمنية ممتدة . فعلى سبيل المثال لتقدير دالة الطلب على البن ، نحتاج إلى بيانات عن كمية البن المطلوبة سنوياً ، على مدى عدة سنوات وليكن من ١٩٦٠ إلى ١٩٨٠ . وبالمثل ، نحتاج إلى بيانات عن متوسط سعر البن ، ودخول المستهلكين ، وسعر الشاى (مثلا ، كبديل للبن) للسنوات من ١٩٦٠ - ١٩٨٠ . وتسمى البيانات عن المتغيرات خلال فترة زمنية ببيانات السلاسل الزمنية . أما تقدير دالة الاستهلاك لمجموعة من العائلات عند نقطة زمنية معينة فإنه يحتاج إلى بيانات مستعرضة *cross — sectional* (بمعنى بيانات عن قيم الإنفاق الاستهلاكى ، والدخل المتاح لكل عائلة فى المجموعة عند نقطة زمنية معينة ولتكن ١٩٨٢) .

١ - ١٣ ماذا يقصد بالآتى : (أ) المرحلة الثالثة لتحليل الاقتصاد القياسى ؟ (ب) المعايير النظرية المسبقة ؟ (ح) المعايير الإحصائية ؟ (د) معايير الاقتصاد القياسى ؟ (هـ) قدرة النموذج على التنبؤ ؟

(أ) المرحلة الثالثة لبحث الاقتصاد القياسى تتضمن تقويم النموذج المقدر على أساس معايير النظرية الاقتصادية المسبقة ، والمعايير الإحصائية ، ومعايير الاقتصاد القياسى وقدرة النموذج على التنبؤ .

(ب) المعايير الاقتصادية المسبقة تشير إلى إشارة وحجم معالم النموذج الذى تفترضه النظرية الاقتصادية . فإذا كانت المعاملات المقدرة لا تتفق وهذه الفروض أو المسلمات ، فإن النموذج يجب أن يعدل أو يرفض .

(ح) المعايير الإحصائية تشير إلى (١) نسبة التغير فى المتغير التابع التى يمكن أن « يشرحها » التغير فى المتغير المستقل أو التفسيرى (٢) التحقق أن مدى تشتت أو انتشار كل معامل مقدر حول القيمة الحقيقية للمعلمة صغير بدرجة كافية لإعطاء الثقة فى التقديرات .

(د) تتعلق معايير الاقتصاد القياسى باختبار ما إذا كانت فروض نموذج الانحدار الأساسى ، وخاصة فيما يتعلق بحد التشويش أو حد الخطأ متحققة .

(هـ) تشير قدرة النموذج على التنبؤ إلى قدرته على التنبؤ بدقة بقيم المتغير التابع باستخدام قيم متوقعة أو معروفة للمتغيرات المستقلة أو التفسيرية .

١ — ١٤ كيف يمكن تقويم دالة الاستهلاك المقدرة فى المعادلة (١ — ٤) من حيث (أ) المعايير النظريه المسبقة ؟ (ب) المعايير الإحصائية ؟ (ﺟ) معايير الاقتصاد القياسى ؟ (د) قدرة النموذج على التنبؤ ؟

(أ) دالة الاستهلاك المقدرة فى المعادلة (١ — ٤) يمكن تقويمها من حيث المعايير النظرية المسبقة بفحص ما إذا كانت المعاملات المقدرة تتفق مع توقعاتها النظرية من حيث الإشارة والحجم كما افترض فى المسألة (١ — ١٠) . فنظرية الطلب كما وردت فى المعادلة (١ — ٤) تتأكد فقط إذا كانت $b_1 < 0$ ، $b_2 > 0$ (إذا كانت X سلعة عادية) ، $b_3 > 0$ (إذا كانت Z سلعة بديلة للسلعة X) ، كما افترض فى نظرية الطلب .

(ب) المعايير الإحصائية تتحقق فقط إذا كانت نسبة « عالية » من التغير فى D_x مع الزمن يمكن « تفسيرها » بالتغيرات فى P_x ، Y و P_z وإذا كان تشتت تقديرات b_1 ، b_2 و b_3 حول المعالم الحقيقية « صغير بدرجة كافية » . وليس هناك إجابة عامة مقبولة لما يعتبر نسبة « عالية » للتغير فى D_x « تفسرها » المتغيرات P_x ، Y و P_z . ومع ذلك فبسبب النزعات المشتركة فى بيانات السلاسل الزمنية يمكن أن نتوقع أن أكثر من 50% إلى 70% من التغير فى المتغير التابع لابد أن يجد تفسيره فى المتغيرات المستقلة إذا حكمنا على النموذج بأنه مرض . وبالمثل ، لحكم على المعاملات المقدرة بأنها « معنوية إحصائياً » ، فأننا نتوقع أن يكون تشتت كل معامل مقدر حول المعلمة الحقيقية (مقيساً بانحرافه المعيارى ، أنظر قسم ٢ — ٣) أقل عموماً من نصف القيمة المقدرة للمعامل .

(ﺟ) تستخدم معايير الاقتصاد القياسى لتحديد ما إذا كانت فروض طرق الاقتصاد القياسى المستخدمة متحققة فى تقدير دالة الطلب فى المعادلة (١ — ٤) . وفى هذه الحالة فقط فإن المعاملات المقدرة يتوفر لها المواصفات المرغوبة من حيث عدم التحيز والاتساق ، والكفاءة ، وهكذا (أنظر قسم ٦ — ٤)

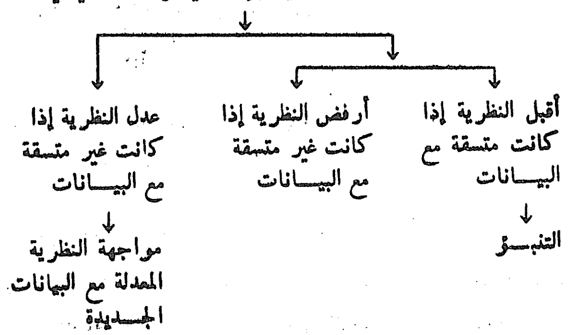
(د) من طرق اختبار مقدرة نموذج الطلب فى المعادلة (١ — ٤) على التنبؤ ، استخدام الدالة المقدرة للتنبؤ بقيمة D_x لفترة لاتشملها العينة وفحص ما إذا كانت القيم المقدرة « قريبة بدرجة كافية » من القيم الفعلية المشاهدة للمتغير D_x خلال هذه الفترة .

١ — ١٥ استخدم رسماً تخطيطياً للتعبير عن المراحل المختلفة لبحوث الاقتصاد القياسي .

المرحلة الأولى : النظرية الاقتصـــادية
↓
النمـــوذج الرياضى
↓
نموذج الاقتصاد القياسى (احتمالى)

المرحلة الثانية : جميع البيانات الملائمة
↓
تقدير معالم النموذج

المرحلة الثالثة : تقويم النموذج على أسس من معايير اقتصادية وإحصائية واقتصاد قياسية
↓



مسائل اضافية

طبيعة عـــلم الإحصـــاء :

١ – ١٦ (أ) فى أى المجالات يعتبر التحليل الإحصائى مهماً ؟ (ب) ماهى أهم وظائف الإحصاء الوصفى ؟ (جـ) ماهى أهم وظائف الإحصاء الاستدلالى ؟

الإجابة : (أ) فى مجالات الاقتصاد والأعمال وغيرها من العلوم الاجتماعية والعلوم الطبيعية .
(ب) تلخيص وتوصيف مجموعات البيانات
(جـ) استنتاج خصائص لمجتمع مامن واقع الخصائص المناظرة لعينة مأخوذة من هذا المجتمع .

١ – ١٧ (أ) هل الاستدلال الإحصائى مرتبط بمنطق استنباطى أو استقرائى ؟ (ب) ماهى شروط صحة الاستدلال الإحصائى ؟
الإجابة : (أ) يرتبط الاستدلال الإحصائى بمنطق استقرائى . (ب) أن يكون من خلال عينة ممثلة ونظرية الاحتمال .

الإحصاء والاقتصاد القياسى :

١ – ١٨ عبر فى معادلة خطية صريحة عن التقرير القائل بأن مستوى الإنفاق الاستثمارى I يرتبط عكسياً مع معدل الفائدة R
الإجابة : (١ – ٥) $I_0 = b_0 + b_1R$ مع افتراض أن b_1 سالبة

١ – ١٩ ماذا تمثل الإجابة على مسألة ١ – ١٨ ؟
الإجابة : تمثل نظرية اقتصادية معبراً عنها فى شكل رياضى دقيق أو محدد

١ – ٢٠ عبر عن المعادلة (١ – ٥) فى صورة احتمالية
الإجابة : (١ – ٦) $I = b_0 + b_1R + u$

١ – ٢١ لماذا يتطلب تحليل الاقتصاد القياسى استخدام صيغ احتمالية ؟
الإجابة : لأن العلاقات بين المتغيرات الاقتصادية غير دقيقة وإلى حد ما شاذة على عكس العلاقات المحددة والدقيقة التى تفترضها النظرية الاقتصادية والاقتصاد الرياضى .

منهاج الاقتصاد القياسى :

١ – ٢٢ ماهى المراحل (أ) الأولى (ب) الثانية (جـ) الثالثة للاقتصاد القياسى ؟
الإجابة : (أ) وضع النظرية فى صيغة معادلة احتمالية مع توضيح الإشارات والقيم المتوقعة للمعالم المقدرة (ب) جميع البيانات عن متغيرات النموذج وتقدير معاملات الدالة (جـ) تقويم المعاملات المقدرة على أسس من الاقتصاد والإحصاء والاقتصاد القياسى .

١ – ٢٣ ماهى المرحلة الأولى فى التحليل الاقتصادى القياسى لنظرية الاستثمار الواردة فى مسألة (١ – ١٨) ؟
الإجابة : وضع النظرية فى صيغة المعادلة (١ – ٦) والتنبؤ بأن $b_1 < 0$

١ – ٢٤ ماهى المرحلة الثانية فى التحليل الاقتصادى القياسى لنظرية الاستثمار الواردة فى مسألة (١ – ١٨) ؟
الإجابة : جمع بيانات سلسلة زمنية عن I و R وتقدير المعادلة (١ – ٦) .

١ – ٢٥ ماهى المرحلة الثالثة فى التحليل الاقتصادى القياسى لنظرية الاستثمار فى مسألة (١ – ١٨) ؟
الإجابة : تحديد ما إذا كان المعامل المقدر $b_1 < 0$ وأن نسبة « كافية » من التغير فى I مع الزمن « يفسره » التغير فى R ، وأن b_1 « معنوية إحصائياً عند المستويات المعتادة » ، وأن فروض الاقتصاد القياسى متحققة

الفصل الثانى

الاحصاء الوصفى

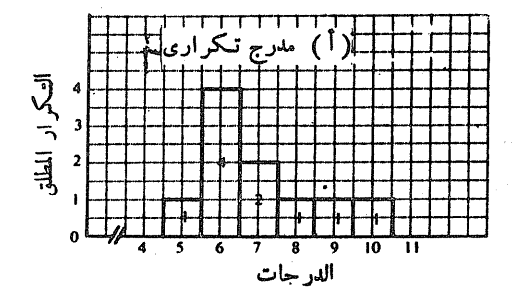
٢ — ١ التوزيعات التكرارية

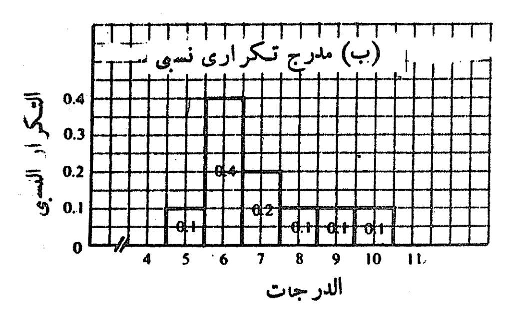
عادة يكون من المفيد تنظيم مجموعة البيانات على شكل توزيع تكرارى . ويكون ذلك بتقسيم البيانات إلى مجموعات أو فئات وتحديد عدد المشاهدات فى كل فئة . ويكون عدد الفئات عادة بين 5 و 15 ويمكن إيجاد **التوزيع التكرارى النسبى** بقسمة عدد المشاهدات فى كل فئة على العدد الإجمالى للمشاهدات ، وبذلك فإن مجموع التكرارات النسبية يساوى 1 . ويمثل التوزيع التكرارى بيانياً باستخدام « **المدرج التكرارى** » حيث تمثل الفئات على المحور الأفقى وتمثل التكرارات على المحور الرأسى أما **المضلع التكرارى** فهو تمثيل بيانى خطى للتوزيع التكرارى ناتج عن توصيل النقاط التى إحداثياتها منتصف الفئة والتكرار ببعضها البعض ويوضح **التوزيع التكرارى المتجمع** بالنسبة لكل فئة إجمالى عدد المشاهدات لجميع الفئات التى تسبق وتشمل هذه الفئة وبرسم التوزيع التكرارى المتجمع نحصل على**منحنى التوزيع أو المنحنى التكرارى المتجمع** .

مثال ١ — حصل طالب على الدرجات الآتية (النهاية العظمى 10) فى عشرة اختبارات أداها أثناء الفصل الدراسى 6 ، 7 ، 6 ، 8 ، 5 ، 7 ، 6 ، 9 ، 10 ، 6 هذه الدرجات يمكن ترتيبها فى شكل توزيع تكرارى كما فى جدول ٢ — ١ كما يمكن عرضها بيانياً كما فى شكل ٢ — ١

جدول ٢ — ١ التوزيع التكرارى للبيانات

الدرجة	التكرار المطلق	التكرار النسبى
5	1	0.1
6	4	0.4
7	2	0.2
8	1	0.1
9	1	0.1
10	1	0.1
	10	1.0





شكل ٢ — ١

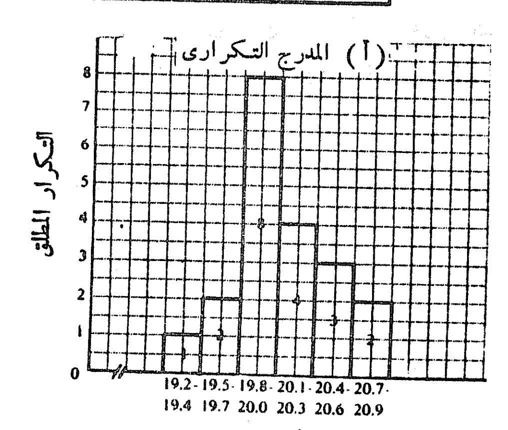
مثال ٢ — تحتوى عينة مكونة من 20 علبة من معلبات الفواكه المحفوظة على وزن صاف يتراوح من 19.3 إلى 20.9 أوقية كما فى جدول ٢–٢ . فإذا أردنا تجميع هذه البيانات فى 6 فئات فإن طول الفئة يكون 0.3 أوقية (19.2 — 21.0)/6 = 0.3 أوقية) ويمكن ترتيب بيانات جدول ٢ — ٢ فى شكل جدول تكرارى كما فى جدول ٢ — ٣ وبيانياً كما فى شكل ٢ — ٢ .

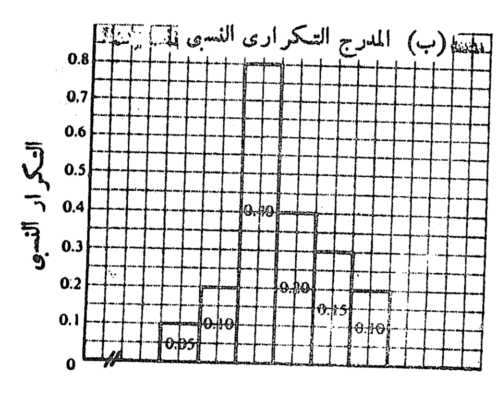
جدول ٢ — ٢ الوزن الصافى للفواكه بالأوقية

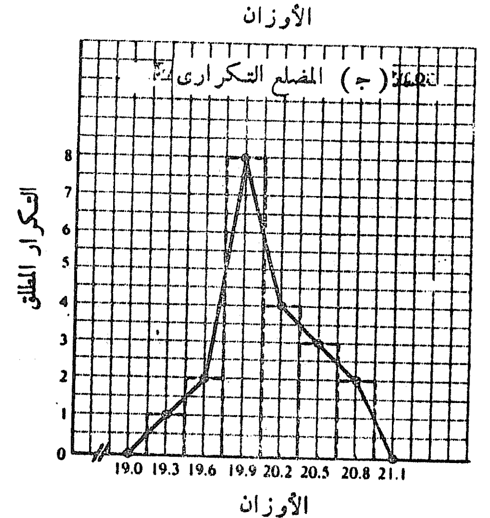
19.7	19.9	20.2	19.9	20.0	20.6	19.3	20.4	19.9	20.3
20.1	19.5	20.9	20.3	20.8	19.9	20.0	20.6	19.9	19.8

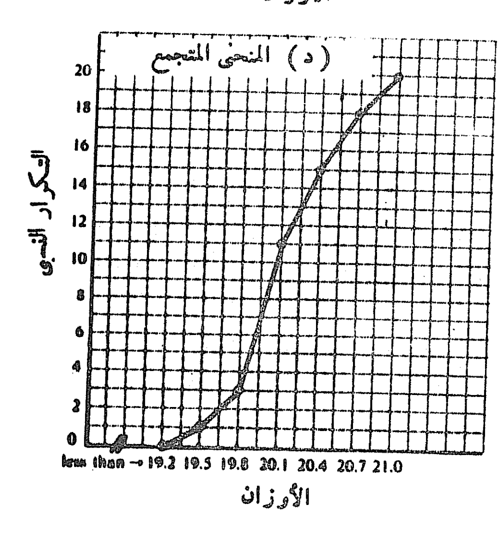
جدول ٢ — ٣ التوزيع التكرارى للأوزان

الوزن بالأوقية	مركز الفئة	التكرار المطلق	التكرار النسبى	التكرار المتجمع
19.2–19.4	19.3	1	0.05	1
19.5–19.7	19.6	2	0.10	3
19.8–20.0	19.9	8	0.40	11
20.1–20.3	20.2	4	0.20	15
20.4–20.6	20.5	3	0.15	18
20.7–20.9	20.8	2	0.10	20
		20	1.00	









شكل ٢ — ٢

٢ـ٢ مقاييس النزعة المركزية

تشير النزعة المركزية إلى موقع التوزيع . وأهم مقاييس النزعة المركزية هى : (١) الوسط الحسابى (٢) الوسيط ، (٣) المنوال . وسوف نقوم بقياس هذه المقاييس بالنسبة المجتمعات (بمعنى مجموعات تشمل جميع العناصر موضع الدراسة) وبالنسبة لعينات مسحوبة من المجتمعات وكذلك بالنسبة البيانات المبوبة والبيانات غير المبوبة .

١ - الوسط الحسابى أو المتوسط ، لمجتمع ما يرمز له بالرمز μ (الحرف اليونانى ميو) ، أما الوسط الحسابى للعينة فيرمز له بالرمز $\bar{X}$ (وتقرأ X bar) . وتحسب μ ، $\bar{X}$ فى حالة البيانات غير المبوبة باستخدام المعادلات الآتية :

(٢ - ١ أ ، ب)
$$\mu = \frac{\sum X}{N} \quad \text{و} \quad \bar{X} = \frac{\sum X}{n}$$

حيث ΣX هى مجموع قيم المشاهدات بينما N ، n تشير إلى عدد المشاهدات فى المجتمع والعينة على الترتيب . أما فى حالة البيانات المبوبة فإن μ ، $\bar{X}$ تحسب كالآتى :

(٢ - ٢ أ ، ب)
$$\mu = \frac{\sum fX}{N} \quad \text{و} \quad \bar{X} = \frac{\sum fX}{n}$$

حيث ΣfX تشير إلى مجموع حاصل ضرب تكرار كل فئة f فى مركز الفئة X .

٢ - الوسيط لبيانات غير مبوبة يشير إلى قيمه المفردة التى تقع فى منتصف المفردات ، بعد ترتيب هذه المفردات تصاعدياً أو تنازلياً . أى أن

(٢ - ٣) الوسيط = الحد الذى ترتيبه $\left(\frac{N+1}{2}\right)$ فى مجموعة البيانات .

حيث N عدد المفردات فى المجتمع (ويستخدم n بدلا منها فى حالة العينة) . وتحسب قيمة الوسيط من بيانات مبوبة كالآتى :

(٢ - ٤) الوسيط = $L + \frac{n/2 - F}{f_m} c$

حيث L = الحد الأدنى للفئة الوسيطية (أى الفئة التى تضم المفردة الوسطى للتوزيع) .

n = عدد المفردات فى مجموعة البيانات

F = مجموع التكرارات فى الفئات السابقة على الفئة الوسيطية .

f_m = تكرار الفئة الوسيطية .

c = طول الفئة

٣ - المنوال هو القيمة الأكثر تكرارا فى مجموعة البيانات . أما بالنسبة للبيانات المبوبة فإن المنوال يحسب كالآتى :

(٢ - ٥) المنوال = $L + \frac{d_1}{d_1 + d_2} c$

حيث L = الحد الأدنى للفئة المنوالية (أى الفئة ذات أكبر تكرار) .

d_1 = الفرق بين تكرار الفئة المنوالية وتكرار الفئة السابقة عليها

d_2 = الفرق بين تكرار الفئة المنوالية والفئة اللاحقة عليها .

c = طول الفئة .

يعتبر الوسط الحسابى أكثر مقاييس النزعة المركزية شيوعاً فى الاستخدام ، غير أن الوسط الحسابى يتأثر بالقيمة المتطرفة ، بينما

لا يتأثر الوسيط أو المنوال بها . وهناك أيضاً كمقاييس للنزعة المركزية ، الوسط الحسابى المرجح ، الوسط الهندسى ، والوسط التوافق ، (انظر مسائل ٢ – ٧ إلى ٢ – ٩) .

مثال٣ – الوسط الحسابى لدرجات 10 امتحانات المعطاة فى مثال ١ يحسب باستخدام معادلة الوسط الحسابى لبيانات غير مبوبة كالآتى :

$$\mu = \frac{\sum X}{N} = \frac{6+7+6+8+5+7+6+9+10+6}{10} = \frac{70}{10} = 7 \text{ درجات}$$

ولإيجاد الوسيط لهذه البيانات غير المبوبة فإننا نرتب أولا القيم العشر للدرجات تصاعدياً : 5, 6, 6, 6, 6, 7, 7, 8, 9, 10 . ثم نحدد الدرجة التى ترتيبها $(N+1)/2$ أو $(10+1)/2 = 5.5$. أى أن الوسيط هو متوسط قيمتى المفردة الخامسة والمفردة السادسة أو $(6+7)/2 = 6.5$. أما المنوال لهذه المجموعة غير المبوبة من البيانات فهو 6 (القيمة التى تكررت أكثر من غيرها فى مجموعة البيانات) .

مثال ٤ – يمكن تقدير الوسط الحسابى للبيانات المبوبة المعطاة فى جدول ٢ – ٣ بمساعدة جدول ٢ – ٤ .

$$\bar{X} = \frac{\sum fX}{n} = \frac{401.6}{20} = 20.8 \text{ أوقية}$$

ويمكن تبسيط الحسابات باستخدام الترميز (انظر مسألة ٢ – ٦) .

جدول ٢ – ٤ حساب الوسط الحسابى لبيانات العينة فى جدول ٢ – ٣

الوزن بالأوقية	مركز الفئة X	التكرار f	fX
19.2–19.4	19.3	1	19.3
19.5–19.7	19.6	2	39.2
19.8–20.0	19.9	8	159.2
20.1–20.3	20.2	4	80.8
20.4–20.6	20.5	3	61.5
20.7–20.9	20.8	2	41.6
		$\sum f = n = 20$	$\sum fX = 401.6$

ويمكن تقدير الوسيط لنفس البيانات المبوبة كالآتى :

$$\text{الوسيط} = L + \frac{n/2 - F}{f_m} c = 19.8 + \frac{20/2 - 3}{8} 0.3 = 19.8 + \frac{7}{8} 0.3$$
$$= 19.8 + 0.2625 \cong 20.06 \text{ oz}$$

حيث $L = 19.8 =$ الحد الأدنى للفئة الوسيطية (أى الفئة 20.0 — 19.8 والتى تحتوى على المشاهدات العاشرة والحادية عشرة) .

$n = 20 =$ عدد المشاهدات أو العناصر .

$F = 3 =$ مجموع التكرارات فى الفئات السابقة على الفئة الوسيطية .

$f_m = 8 =$ تكرار الفئة الوسيطية .

$c = 0.3 =$ طول الفئة .

وبالمثل ،

$$\text{المنوال} = L + \frac{d_1}{d_1 + d_2} c = 19.8 + \frac{6}{6+4} 0.3 = 19.8 + \frac{1.8}{10} = 19.8 + 0.18 = 19.98 \text{ أوقية}$$

وكما ذكر فى المسألة ٢ — ٤ ، فإن الوسط الحسابى والوسيط والمنوال للبيانات المبوبة تعتبر تقديرات تستخدم فقط عندما تكون البيانات المبوبة متاحة أو عندما يراد اختصار الحسابات لمجموعة كبيرة من البيانات غير المبوبة .

٢—٣ مقاييس التشتت

يشير **التشتت** إلى اختلاف أو انتشار البيانات . وأهم مقاييس التشتت هى (١) **الانحراف المتوسط** ، (٢) **التباين** ، (٣) **الانحراف المعيارى** ، وسنقوم بقياس هذه المقاييس بالنسبة للمجتمعات والعينات وكذلك بالنسبة للبيانات المبوبة والبيانات غير المبوبة .

١ — **الانحراف المتوسط** (AD) للبيانات غير المبوبة يحسب كالآتى :

$$AD = \frac{\sum |X - \mu|}{N}$$ للمجتمعات (٢ — ٦ أ)

$$AD = \frac{\sum |X - \bar{X}|}{n}$$ للعينات (٢ — ٦ ب)

حيث يشير الخطان الرأسيان إلى استخدام القيمة المطلقة للانحراف ، أى القيمة الموجبة للانحراف ، مع بقاء باقى الرموز بنفس المعنى المستخدم فى قسم ٢ — ٢ . وبالنسبة للبيانات المبوبة

$$AD = \frac{\sum f|X - \mu|}{N}$$ للمجتمعات (٢ — ٧ أ)

$$AD = \frac{\sum f|X - \bar{X}|}{n}$$ للعينات (٢ — ٧ ب)

حيث f ترمز إلى تكرار الفئة و X ترمز إلى مركز الفئة .

٢ — **التباين** . تباين المجتمع σ^2 (الحرف اليونانى سيجما تربيع) وتباين العينة s^2 للبيانات غير المبوبة تحسب كالآتى :

(٢ — ٨ أ ، ب)

$$\sigma^2 = \frac{\sum (X - \mu)^2}{N} \quad و \quad s^2 = \frac{\sum (X - \bar{X})^2}{n - 1}$$

وللبيانات المبوبة ،

$$\sigma^2 = \frac{\sum f(X - \mu)^2}{N} \quad و \quad s^2 = \frac{\sum f(X - \bar{X})^2}{n - 1}$$ (٢ — ٩ أ ، ب)

٣ — **الانحراف المعيارى** . الانحراف المعيارى للمجتمع σ وللعينة s هما الجذر التربيعى الموجب للتباين المناظر لكل منهما . فبالنسبة للبيانات غير المبوبة ،

$$\sigma = \sqrt{\frac{\sum (X - \mu)^2}{N}} \quad و \quad s = \sqrt{\frac{\sum (X - \bar{X})^2}{n - 1}}$$ (٢ — ١٠ أ ، ب)

وبالنسبة للبيانات المبوبة ،

$$\sigma = \sqrt{\frac{\sum f(X - \mu)^2}{N}} \quad و \quad s = \sqrt{\frac{\sum f(X - \bar{X})^2}{n - 1}}$$ (٢ — ١١ أ ، ب)

ويعتبر الانحراف المعيارى أكثر مقاييس التشتت المطلق شيوعاً فى الاستخدام . وهناك مقاييس أخرى (بجانب التباين والانحراف المتوسط) وهى المدى ، المدى الربيعى ، والانحراف الربيعى (انظر مسائل ٢ — ١١ و ٢ — ١٢) .

٤ – معامل الاختلاف V يقيس التشتت النسبى :

(٢ – ١٢ أ) للمجتمعات $V = \frac{\sigma}{\mu}$

(٢ – ١٢ ب) للعينات $V = \frac{s}{\bar{X}}$

مثال ٥ – يمكن إيجاد الانحراف المتوسط والتباين والانحراف المعيارى ومعامل الاختلاف للبيانات غير المبوبة فى مثال ١ بالاستعانة بجدول ٢ – ٥ ($\mu = 7$ ، انظر مثال ٣) :

$$\text{AD} = \frac{\sum|X-\mu|}{N} = \frac{12}{10} = 1.2 \quad \text{درجات}$$

$$\sigma^2 = \frac{\sum(X-\mu)^2}{N} = \frac{22}{10} = 2.2 \quad \text{درجات مربعة}$$

$$\sigma = \sqrt{\frac{\sum(X-\mu)^2}{N}} = \sqrt{\frac{22}{10}} = \sqrt{2.2} \cong 1.48 \quad \text{درجات}$$

$$V = \frac{\sigma}{\mu} \cong \frac{1.48}{7} \cong 0.21, \quad \text{أو} \quad 21\%$$

جدول ٢ – ٥ عمليات حسابية على بيانات المثال ١

الدرجة	μ	$X-\mu$	$\lvert X-\mu\rvert$	$(X-\mu)^2$
6	7	−1	1	1
7	7	0	0	0
6	7	−1	1	1
8	7	1	1	1
5	7	−2	2	4
7	7	0	0	0
6	7	−1	1	1
9	7	2	2	4
10	7	3	3	9
6	7	−1	1	1
		$\sum(X-\mu)=0$	$\sum\lvert X-\mu\rvert=12$	$\sum(X-\mu)^2=22$

مثال ٦ – يمكن حساب الانحراف المتوسط والتباين ، والانحراف المعيارى ، ومعامل الاختلاف للتوزيع التكرارى للأوزان (بيانات مبوبة) الواردة بجدول ٢ – ٣ بالاستعانة بجدول ٢ – ٦ ($\bar{X} = 20.08$ أوقية ، انظر مثال ٤) :

$$\text{AD} = \frac{\sum f|X-\bar{X}|}{n} = \frac{6.36}{20} = 0.318 \quad \text{أوقية}$$

$$s^2 = \frac{\sum f(X-\bar{X})^2}{n-1} = \frac{2.9520}{19} \cong 0.1554 \quad \text{أوقية مربعة}$$

$$s = \sqrt{\frac{\sum f(X-\bar{X})^2}{n-1}} = \sqrt{\frac{2.9520}{19}} = \sqrt{0.1544} \cong 0.3942 \quad \text{أوقية}$$

$$V = \frac{s}{\bar{X}} \cong \frac{0.3942 \text{ oz}}{20.08 \text{ oz}} \cong 0.0196, \quad \text{أو} \quad 1.96\%$$

لاحظ أنه فى معادلة تقدير كل من s^2 و s ، نستخدم $n - 1$ وليس n فى المقام (انظر مسألة ٢ – ١٦ للتعليل) . ويمكن من المعادلات المستخدمة لتقدير σ^2 ، σ ، s^2 و s اشتقاق معادلات أخرى لتبسيط الحسابات إذا كان حجم البيانات كبيراً (انظر مسائل ٢ – ١٧ إلى ٢ – ١٩ لمعرفة كيفية اشتقاق وتطبيق هذه المعادلات) .

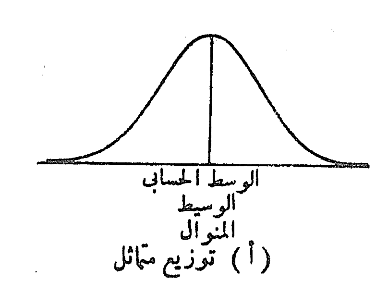
جدول ٢ – ٦ عمليات حسابية على بيانات جدول ٢ – ٤

الوسط الحسابي	التكرار	الوزن بالأوقية	مركز الفئة	$X - \bar{X}$	$\lvert X - \bar{X}\rvert$	$\sum f\lvert X - \bar{X}\rvert$	$(X - \bar{X})^2$	$f(X - \bar{X})^2$
19.20–19.40	19.30	1	20.08	− 0.78	0.78	0.78	0.6084	0.6084
19.50–19.70	19.60	2	20.08	− 0.48	0.48	0.96	0.2304	0.4608
19.80–20.00	19.90	8	20.08	− 0.18	0.18	1.44	0.0324	0.2592
20.10–20.30	20.20	4	20.08	0.12	0.12	0.48	0.0144	0.0576
20.40–20.60	20.50	3	20.08	0.42	0.42	1.26	0.1764	0.5292
20.70–20.90	20.80	2	20.08	0.72	0.72	1.44	0.5184	1.0368
		$\sum f = n = 20$				$\sum f\lvert X - \bar{X}\rvert = 6.36$		$\sum f(X - \bar{X})^2 = 2.9520$

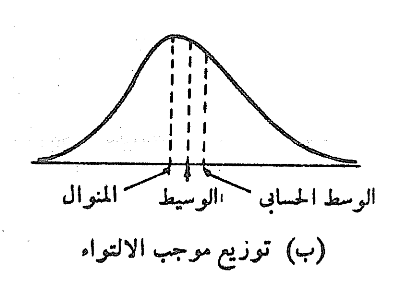
٢–٤ اشكال التوزيعات التكرارية

يشير شكل التوزيع التكرارى إلى (١) تماثل التوزيع من عدمه (التواء التوزيع) (٢) تدبب التوزيع (تفرطح) .

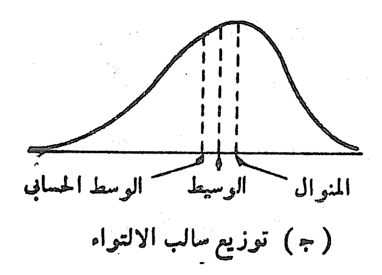
١ – الالتواء . يكون التواء التوزيع صفراً إذا كان التوزيع متماثلا حول الوسط الحسابى . وفى حالة التوزيع المتماثل (ذى المنوال الواحد) فإن الوسط الحسابى يساوى الوسيط يساوى المنوال . والتوزيع موجب الالتواء هو التوزيع الذى يكون طرفه الأيمن أطول . وعندئذ يكون الوسط الحسابى > الوسيط > المنوال . ويكون التوزيع سالب الالتواء هو التوزيع الذى يكون طرفه الأيسر أطول . وعندئذ يكون المنوال > الوسيط > الوسط الحسابى (انظر شكل ٢ – ٣) .



(أ) توزيع متماثل



(ب) توزيع موجب الالتواء



(ﺟ) توزيع سالب الالتواء

شكل ٢ – ٣

ويمكن قياس الالتواء باستخدام معامل بيرسون للالتواء :

$$\text{Sk} = \frac{3(\mu - \text{med})}{\sigma}$$ للمجتمعات (٢ – ١٣ أ)

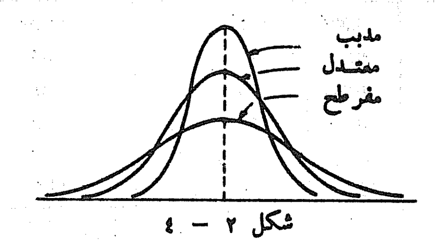
$$\text{Sk} = \frac{3(\bar{X} - \text{med})}{s}$$ للعينات (٢ – ١٣ ب)

والالتواء أيضاً يمكن قياسه بالعزم الثالث (بسط المعادلة ٢ – ١٤ أ ، ب) مقسوماً على مكعب الانحراف المعيارى :

$$\text{Sk} = \frac{\sum f(X - \mu)^3}{N\sigma^3}$$ للمجتمعات (٢ – ١٤ أ)

$$\text{Sk} = \frac{\sum f(X - \bar{X})^3}{n s^3}$$ للعينات (٢ – ١٤ ب)

٢ – التفرطح . التوزيع ذو القمة العالية يسمى مدبباً ، وعلى العكس يسمى التوزيع ذو القمة المنبسطة مفرطحاً وذلك بالقياس إلى المعتدل أو متوسط التفرطح (انظر شكل ٢ – ٤) . ويمكن لقياس التفرطح استخدام العزم الرابع (البسط فى المعادلة (٢ – ١٥ أ ، ب)) مقسوماً على الانحراف المعيارى مرفوعاً للقوة الرابعة . علماً بأن معامل التفرطح للتوزيع المعتدل = 3



شكل ٢ – ٤

$$\text{معامل التفرطح} = \frac{\sum f(X-\mu)^4}{N\sigma^4}$$ للمجتمعات (٢ – ١٥ أ)

$$\text{معامل التفرطح} = \frac{\sum f(X-\overline{X})^4}{n\,s^4}$$ للعينات (٢ – ١٥ ب)

مثال ٧ – يمكن تقدير معامل بيرسون للالتواء للدرجات فى مثال ١ باستخدام $\mu = 7$ ، $med = 6.5$ (انظر مثال ٣)، و $\sigma = 1.48$ (انظر مثال ٥) :

$$Sk = \frac{3(\mu - med)}{\sigma} \cong \frac{3(7-6.5)}{1.48} \cong \frac{3(0.5)}{1.48} \cong 1.01$$ (انظر شكل ٢ – ١)

وبالمثل ، باستخدام $\overline{X} = 20.08$ oz ، الوسيط = 20.06 oz (انظر مثال ٤) ، و $s = 0.39$ oz (أنظر مثال ٦) ، يمكن تقدير معامل بيرسون للالتواء للتوزيع التكرارى للأوزان فى جدول ٢ – ٣ كما يأتى :

$$Sk = \frac{3(\overline{X} - med)}{s} \cong \frac{3(20.08-20.06)}{0.39} \cong 0.15$$ (انظر شكل ٢ – ٢ جـ)

بالنسبة للتفرطح ، انظر مسألة ٢ – ٢٣ .

مسائل محلولة

التوزيعات التكرارية :

٢ – ١ جدول ٢ – ٧ يبين درجات اختبار ما لفصل من 40 طالباً . (أ) رتب هذه الدرجات (مجموعة البيانات الخام) فى جدول يبدأ بأصغر الدرجات وينتهى بأكبرها . (ب) كون جدولا موضحاً أطوال الفئات ومراكز الفئات والتكرار المطلق والنسبى والمتجمع لكل درجة (جـ) اعرض البيانات فى شكل مدرج تكرارى ، مدرج تكرارى نسبى ، مضلع تكرارى ،ومنحنى متجمع صاعد .

جدول ٢ – ٧ درجات الاختبار لفصل من 40 طالبا

7	5	6	2	8	7	6	7	3	9
10	4	5	5	4	6	7	4	8	2
3	5	6	7	9	8	2	4	7	9
4	6	7	8	3	6	7	9	10	5

(أ) انظر جدول ٢ – ٨

جدول ٢ – ٨ بيان الدرجات مرتبة تصاعدياً

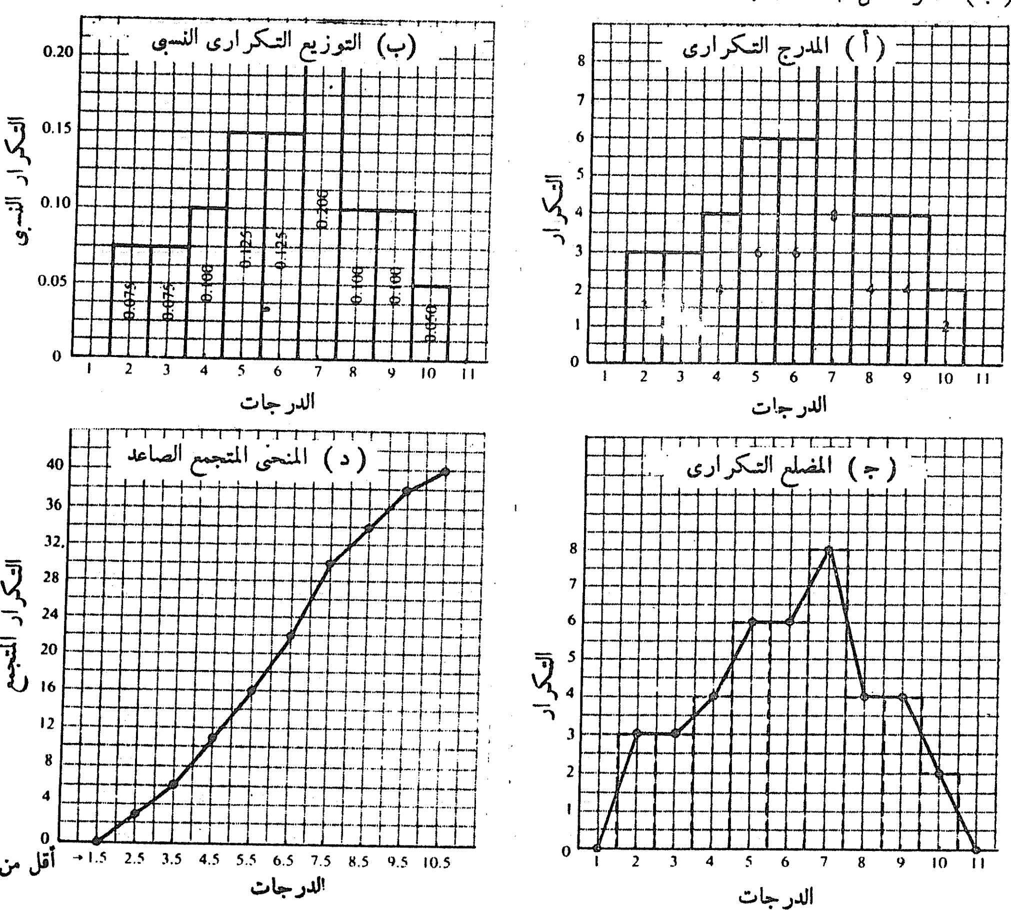
2	2	2	3	3	3	4	4	4	4
4	5	5	5	5	5	6	6	6	6
6	6	7	7	7	7	7	7	7	7
8	8	8	8	9	9	9	9	10	10

(ب) انظر جدول ٢ – ٩ . لاحظ أنه طالما أننا نتعامل هنا مع بيانات منفصلة (أى معبرا عنها باستخدام أعداد صحيحة) ، فقد استخدمنا الدرجات الفعلية كمراكز للفئات .

جـدول ٢ – ٩ توزيع تكرارى للدرجات

الدرجة	مركز الفئة	التكرار المطلق	التكرار النسبى	التكرار المتجمع
1.5–2.4	2	3	0.075	3
2.5–3.4	3	3	0.075	6
3.5–4.4	4	5	0.125	11
4.5–5.4	5	5	0.125	16
5.5–6.4	6	6	0.150	22
6.5–7.4	7	8	0.200	30
7.5–8.4	8	4	0.100	34
8.5–9.4	9	4	0.100	38
9.5–10.4	10	2	0.050	40
		40	1.000	

(جـ) انظر شكل ٢ – ٥ .



شكل ٢ – ٥

٢ - ٢ إذا كان أجر الساعة لعينة مكونة من 25 عاملا بأحد المصانع هو كما فى جدول ٢ - ١٠ .

(أ) رتب هذه البيانات الخام فى جدول تبدأ بالأجر الأصغر وتنتهى بالأجر الأعلى .

(ب) جمع هذه البيانات فى فئات .

(ج) أعرض البيانات فى شكل مجمع تكرارى ، مدرج تكرارى نسبى ، مضلع تكرارى ، ومنحنى متجمع صاعد .

جدول ٢ - ١٠ أجر الساعة بالدولار

3.65	3.78	3.85	3.95	4.00	4.10	4.25	3.55	3.85	3.96
3.60	3.90	4.26	3.75	3.95	4.05	4.08	4.15	3.80	4.05
3.88	3.95	4.06	4.18	4.05					

(أ) انظر جدول ٢ - ١١ .

جدول ٢ - ١١ بيان الأجور بالدولار مرتبة تصاعدياً .

3.55	3.60	3.65	3.75	3.78	3.80	3.85	3.85	3.88	3.90
3.95	3.95	3.95	3.96	4.00	4.05	4.05	4.05	4.06	4.08
4.10	4.15	4.18	4.25	4.26					

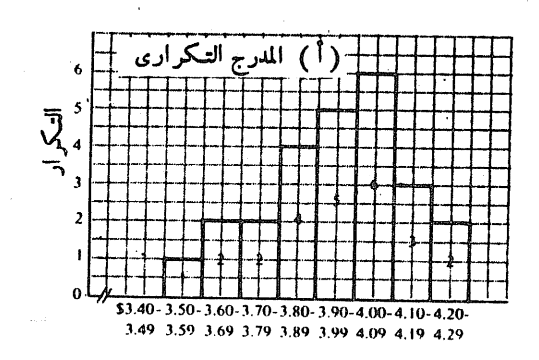
(ب) من جدول ٢ - ١٠ يلاحظ أن أقل أجر للساعة 3.55 $ وأعلى أجر 4.26 $. ويمكن تقسيم هذا المدى إلى 8 فئات متساوية طول كل منها 0.10 $. أى ، 0.10 $ = 0.80/8 $ = ($ 4.30 — $ 3.50)/8 لاحظ أنه قد تم توسيع المدى إلى « من 3.50 $ إلى 4.30 $ » وذلك حتى يقع أقل أجر 3.55 $ داخل الفئة الأولى ،ويقع أعلى أجر ، 4.26 $ داخل الفئة الأخيرة . ومن الملائم أيضاً إيجاد مركز كل فئة (وسوف نحتاج هذا لرسم المضلع التكرارى) . وتوضح هذه فى الجدول ٢ - ١٢ .

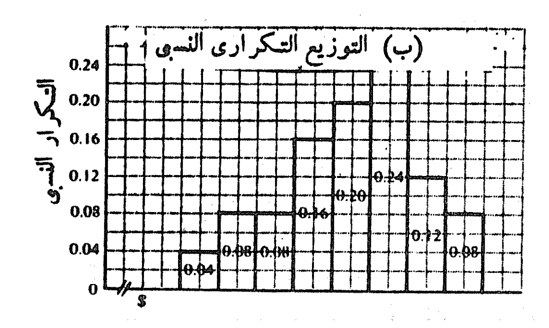
جدول ٢ - ١٢ التوزيع التكرارى للأجور

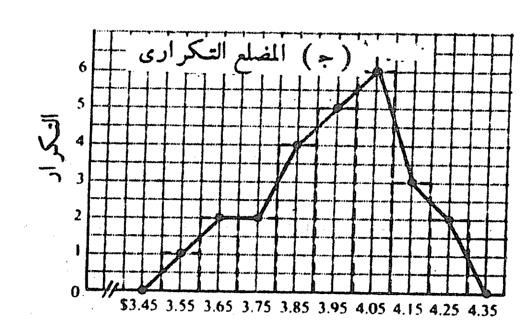
أجر الساعة	مركز الفئة	التكرار المطلق	التكرار النسبى	التكرار المتجمع
$3.50–3.59	$3.55	1	0.04	1
3.60–3.69	3.65	2	0.08	3
3.70–3.79	3.75	2	0.08	5
3.80–3.89	3.85	4	0.16	9
3.90–3.99	3.95	5	0.20	14
4.00–4.09	4.05	6	0.24	20
4.10–4.19	4.15	3	0.12	23
4.20–4.29	4.25	2	0.08	25
		25	1.00	

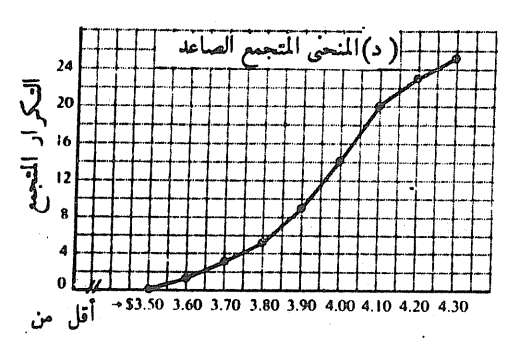
(ج) انظر شكل ٢ - ٦ . ويمكن الحصول على المنحنى المتجمع الصاعد برسم التكرار المتجمع عند القيم 3.595 $ ، 3695 $ ، 3.795 $ وهكذا (وذلك حتى تشمل الحد الأعلى لكل فئة) وعادة يستخدم تعبير حدود الفئات أو الحدود الدقيقة للاشارة إلى القيم 3.595 $ ، 3.695 $ ، 3.795 $ إلى آخره . لاحظ أن مركز الفئة يتم الحصول عليه باضافةالحد الأدنى والحد الأعلى للفئة والقسمة على 2 . فمثلا ، مركز الفئة الثانية يساوى

$(3.595 + 3.695)/2 = 7.290/2 = 3.65$ (انظر جدول ٢ - ١٢)









شكل ٢ – ٦

مقاييس النزعة المركزية :

٢ – ٣ أوجد الوسط الحسابى والوسيط والمنوال (أ) لدرجات اختبار الفصل المكون من 40 طالباً المعطاة فى جدول ٢ – ٧ (بيانات غير مبوبة) ، (ب) للبيانات المبوبة لهذه الدرجات المعطاة فى جدول ٢ – ٩ .

(أ) حيث أننا نتعامل هنا مع كل الدرجات ، فنحن نريد حساب الوسط الحسابى للمجتمع :

$$\mu = \frac{\sum X}{N} = \frac{7+5+6+\cdots+5}{40} = \frac{240}{40} = 6 \text{ درجات}$$

أى أن μ يمكن الحصول عليها بجمع كل الدرجات الأربعين المعطاة فى جدول ٢ – ٧ والقسمة على 40 (وضعت النقاط الثلاث فى وسط الأرقام فى معادلة حساب الوسط الحسابى بعاليه لتجنب سرد كل الأربعين مفردة الواردة فى جدول ٢ – ٧) .

الوسيط هو قيمة العنصر الذى ترتيبه $(N+1)/2$ فى البيانات الواردة فى جدول ٢ – ٨ . ومن ثم ، فإن الوسيط هو قيمة العنصر الذى ترتيبه $(40+1)/2 = 20.5$ ، أى متوسط قيمة المفردتين اللتين تقعان عند ترتيب 20 ، 21 وحيث أن كلا منهما تساوى 6 ، فإن الوسيط يكون 6 . أما **المنوال** فقيمته 7 (القيمة الأكثر تكراراً فى مجموعة البيانات) .

(ب) يمكن إيجاد الوسط الحسابى للمجتمع مع البيانات المبوبة فى جدول ٢ – ٩ مع الاستعانة بجدول ٢ – ١٣ :

$$\mu = \frac{\sum fX}{N} = \frac{240}{40} = 6$$

وهو نفس الوسط الحسابى السابق الحصول عليه من البيانات غير المبوبة . لاحظ أن مجموعة التكرارات Σf يساوى عدد المشاهدات فى المجتمع N وأن $\Sigma X = \Sigma fX$. أما الوسيط من البيانات المبوبة فى جدول ٢ – ١٣ فيكون

$$\text{الوسيط} = L + \frac{N/2 - F}{f_m} c = 5.5 + \frac{40/2 - 16}{6} 1 = 5.5 + 0.67 = 6.17$$

حيث $L = 5.5$ = الحد الأدنى للفئة الوسيطية (أى الفئة من 5.5 إلى 6.4 ، والتى تحتوى على المشاهدات التى ترتيبها 20 و 21)) .

$N = 40$ = عدد المشاهدات

$F = 16$ = مجموع المشاهدات فى الفئات السابقة على الفئة الوسيطية

$f_m = 6$ = تكرار الفئة الوسيطية

$c = 1$ = طول الفئة

ويحسب المنوال للبيانات المبوبة فى جدول ٢ – ١٣ كما يلى :

$$\text{المنوال} = L + \frac{d_1}{d_1 + d_2} c = 6.5 + \frac{2}{2+4} 1 = 6.5 + 0.33 = 6.83$$

حيث $L = 6.5$ = الحد الأدنى للفئة المنوالية (أى الفئة من 6.5 إلى 7.4 والتى يقابلها أعلى تكرار 8) .

$d_1 = 2$ = تكرار الفئة المنوالية 8 ، مطروحاً منه تكرار الفئة قبل المنوالية 6 .

$d_2 = 4$ = تكرار الفئة المنوالية ، 8 ، مطروحاً منه تكرار الفئة بعد المنوالية ، 4

$c = 1$ = طول الفئة

لاحظ أنه بينما يتطابق الوسط الحسابى المحسوب من بيانات مبوبة مع الوسط الحسابى المحسوب من بيانات غير مبوبة فإن كلا من الوسيط والمنوال يعطى فقط تقديراً تقريبياً .

جدول ٢ – ١٣ حساب الوسط الحسابى للمجتمع للبيانات المبوبة فى جدول ٢ – ٩

الدرجة	مركز الفئة X	التكرار f	fX
1.5–2.4	2	3	6
2.5–3.4	3	3	9
3.5–4.4	4	5	20
4.5–5.4	5	5	25
5.5–6.4	6	6	36
6.5–7.4	7	8	56
7.5–8.4	8	4	32
8.5–9.4	9	4	36
9.5–10.4	10	2	20
		$\sum f = N = 40$	$\sum fX = 240$

٢ – ٤ أوجد الوسط الحسابى والوسيط والمنوال (أ) لعينة الأجور لعدد 25 عاملا الواردة فى جدول ٢ – ١٠ (البيانات غير المبوبة) (ب) للبيانات المبوبة لهذه الأجور المعطاة فى جدول ٢ – ١٢ .

(أ) $$\bar{X} = \frac{\sum X}{n} = \frac{\$3.65 + \$3.78 + \$3.85 + \cdots + \$4.05}{25} = \frac{\$98.65}{25} = \$3.946 \text{ أو } \$3.95$$

. الوسيط = 3.95 $ (قيمة المفردة التى ترتيبها 13 = 2/(1 + 25) = 2/(1 + n) فى مجموعة البيانات فى جدول (٢-١١) .

المنوال = 3.95 $ و 4.05 $ ، حيث هناك 3 مفردات تقابل كل قيمة منهما . أى أن التوزيع ذو منوالين .

(ب) يمكن إيجاد الوسط الحسابى للعينة للبيانات المبوبة الواردة فى جدول ٢ – ١٢ مع الاستعانة بجدول ٢ – ١٤ :

$$\bar{X} = \frac{\sum fX}{n} = \frac{\$98.75}{25} = \$3.95$$

لاحظ أنه فى هذه الحالة $\Sigma fX = \$\,98.75 \neq \Sigma X = \$\,98.65$ (كما فى أ) حيث أن متوسط المشاهدات فى كل فئة لا يساوى مركز الفئة لكل الفئات (كما فى مسألة ٢ – ٣ (ب)) . وعليه فإن $\bar{X}$ المحسوبة للبيانات المبوبة تعتبر فقط تقريباً لقيمة $\bar{X}$ الحقيقية المحسوبة من البيانات غير المبوبة . وكثيرا ما تتوافر لدينا البيانات فقط فى صورة مبوبة ، أو قد تتوفر البيانات غير مبوبة فى مجموعة كبيرة جداً بحيث أن تقدير الوسط الحسابى بعد تجميع البيانات فى فئات يوفر كثيراً فى العمل الحسابى .

$$\text{الوسيط} = L + \frac{n/2 - F}{f_m}c = \$3.90 + \frac{25/2 - 9}{5}(0.10) = \$3.90 + \$0.07 = \$3.97$$

بالمقارنة مع قيمة الوسيط الحقيقية 3.95 $ للبيانات غير المبوبة (انظر أ)

$$\text{المنوال} = L + \frac{d_1}{d_1 + d_2}c = \$4.00 + \frac{1}{1+3}(0.10) = \$4.00 + \$0.025 = \$4.025 \quad \text{or} \quad \$4.03$$

بالمقارنة مع القيم الحقيقية للمنوال 3.95 $ و 4.05 $ للبيانات غير المبوبة (انظر أ) . وأحياناً يستخدم مركز الفئة المنوالية كتقريب للمنوال .

جدول ٢ – ١٤ حساب الوسط الحسابى للبيانات المبوبة فى جدول ٢ – ١٢

أجر الساعة	مركز الفئة X	التكرار f	fX
$3.50–3.59	$3.55	1	$3.55
3.60–3.69	3.65	2	7.30
3.70–3.79	3.75	2	7.50
3.80–3.89	3.85	4	15.40
3.90–3.99	3.95	5	19.75
4.00–4.09	4.05	6	24.30
4.10–4.19	4.15	3	12.45
4.20–4.29	4.25	2	8.45
		$\sum f = n = 25$	$\sum fX = \$98.75$

٢ – ٥ أذكر مزايا وعيوب (أ) الوسط الحسابى (ب) الوسيط ، (جـ) المنوال ، كمقاييس للنزعة المركزية .

(أ) مزايا الوسط الحسابى هى (١) أنه مقياس مألوف وسهل الفهم (٢) أنه يأخذ جميع مفردات المجموعة فى الاعتبار (٣) أنه يستخدم فى حساب كثير من المقاييس والاختبارات الإحصائية الأخرى . عيوب الوسط الحسابى هى (١) أنه يتأثر بالقيم المتطرفة (٢) أنه يستغرق وقتاً طويلا للحساب من مجموعة كبيرة من البيانات غير المبوبة (٣) لا يمكن حسابه إذا كان الجدول فى حالة البيانات المبوبة مفتوحاً من أحد طرفيه .

(ب) مزايا الوسيط هى (١) لا يتأثر بالقيم المتطرفة (٢) يمكن فهمه بسهولة (نصف البيانات أصغر من الوسيط والنصف الثانى أكبر منه) (٣) يمكن حسابه فى جداول مفتوحة وأيضاً إذا كانت البيانات كيفية وليس فقط البيانات

الكمية . عيوب الوسيط هى (أ) لا يستخدم الكثير من البيانات المتاحة (ب) نحتاج معه إلى ترتيب المفردات تصاعدياً مما يستغرق وقتاً طويلا إذا كانت مجموعة البيانات كبيرة .

(جـ) مزايا المنوال هى نفس مزايا الوسيط . عيوب المنوال هى (١) كما فى حالة الوسيط ، لا يستخدم المنوال الكثير من البيانات المتاحة . (٢) فى بعض الأحيان لا يوجد منوال حيث لا يتكرر أى من القيم أكثر من مرة ، وفى أحيان أخرى يكون هناك أكثر من منوال . وبصفة عامة فإن الوسط الحسابى يعتبر أكثر مقاييس النزعة المركزية استخداماً ، كما يعتبر المنوال أقلها استخداماً .

٢ – ٦ أوجد الوسط الحسابى للبيانات المبوبة فى جدول ٢ – ١٢ باستخدام الترميز (الطريقة المختصرة ، وذلك بتعيين القيمة $\mu = 0$ للفئة الرابعة أو الخامسة والقيم $\mu = -1$ ، $\mu = -2$ وهكذا للفئات السابقة عليها والقيم $\mu = 1$ ، $\mu = 2$ وهكذا للفئات اللاحقة ثم استخدام الصيغة .

$$\bar{X} = X_0 + \frac{\sum f\mu}{n} c \qquad (٢ - ١٦)$$

حيث X_0 هى مركز الفئة التى عين لها القيمة $\mu = 0$ ، c هى طول الفئات) . انظر جدول ٢ – ١٥ .

جدول ٢ – ١٥ حساب الوسط الحسابى باستخدام الترميز للبيانات المبوبة فى جدول ٢ – ١٢

أجر الساعة	مركز الفئة X	الرمز μ	التكرار f	$f\mu$
$3.50–3.59	$3.55	−3	1	−3
3.60–3.69	3.65	−2	2	−4
3.70–3.79	3.75	−1	2	−2
3.80–3.89	3.85	0	4	0
3.90–3.99	3.95	1	5	5
4.00–4.09	4.05	2	6	12
4.10–4.19	4.15	3	3	9
4.20–4.29	4.25	4	2	8
			$\sum f = n = 25$	$\sum f\mu = 25$

$$\bar{X} = X_0 + \frac{\sum f\mu}{n} c = \$3.85 + \frac{25}{25}(\$0.10) = \$3.85 + \$0.10 = \$3.95$$

ويلاحظ أن الوسط الحسابى $\bar{X}$ المحسوب باستخدام الترميز يتطابق مع ذلك المحسوب بدون استخدامه فى مسألة ٢ – ٤ ب ، بينما يتخلص الترميز من مشاكل التعامل مع القيم الكبيرة لمراكز الفئات ، ومن ثم فإنه يساهم فى تبسيط العمليات الحسابية .

٢ – ٧ شركة تدفع أجراً قدره 4 $ فى الساعة لعمالها غير المهرة وعددهم 25 و 6 $ فى الساعة للعمال شبه المهرة وعددهم 15 ، و 8 $ للعمال المهرة وعددهم 10. ما هو المتوسط المرجح أو الوسط الحسابى المرجح للأجور التى تدفعها الشركة ؟

لإيجاد الوسط الحسابى المرجح ، أو المتوسط المرجح ، للمجتمع μ_w أو للعينة $\bar{X}_w$ ، فإن الأوزان w لها نفس وضع التكرار عند إيجاد الوسط الحسابى من بيانات مبوبة . أى أن

$$\mu_w = \frac{\sum wX}{\sum w} \quad \text{أو} \quad \bar{X}_w \qquad (٢ - ١٧)$$

وبالنسبة لهذه المسألة فإن الأوزان هى عدد العمال المقابلة لكل أجر ، ومجموع الأوزان Σw يساوى مجموع العمال :

$$\mu_w = \frac{(\$4)(25) + (\$6)(15) + (\$8)(10)}{25 + 15 + 10} = \frac{\$100 + \$90 + \$80}{50} = \frac{\$270}{50} = \$5.40$$

بينما أن المتوسط البسيط [$\$6 = (\$4 + \$6 + \$8)/3$] $6 . يمكن القول أن المتوسط المرجح هو مقياس أفضل لمتوسط الأجور .

٢ – ٨ إذا كان معدل التضخم لشعب ما هو 2% فى السنة الأولى ، 5% فى السنة الثانية ، 12.5% فى السنة الثالثة ، أوجد الوسط الهندسى لمعدلات التضخم (الوسط الهندسى μ_G أو $\bar{X}_G$ لمجموعة موجبة من الأرقام n هو الجذر النونى لحاصل ضرب هذه الأرقام ويستخدم أساساً لإيجاد متوسط لمعدلات التغير وللأرقام القياسية) . أى

$$\mu_G \quad \text{or} \quad \bar{X}_G = \sqrt[n]{X_1 \cdot X_2 \cdots X_n} \qquad (٢ - ١٨)$$

حيث X_1 ، X_2 ، .... X_n تشير إلى عدد n (أو N) من المشاهدات .

$$\mu_G = \sqrt[3]{(2)(5)(12.5)} = \sqrt[3]{125} = 5\%$$

بينما أن الوسط الحسابى البسيط $\mu = (2 + 5 + 12.5)/3 = 19.5 / 3 = 6.5\%$ وإذا تساوت قيم جميع المفردات فإن μ_G تساوى μ ، وفى غير ذلك فإن μ_G تكون أصغر من μ . وعادة يتم حساب μ_G باستخدام اللوغاريتمات :

$$\log \mu_G = \frac{\sum \log \mu}{N} \qquad (٢ - ١٩)$$

ويستخدم الوسط الهندسى أساساً فى رياضيات التمويل والإدارة المالية .

٢ – ٩ تقطع مسافرة مسافة قدرها 10 mi على الطريق خارج المدينة بسرعة قدرها 60 mi/h ومسافة قدرها 10 mi على طرق داخل المدينة بسرعة قدرها 15 mi / h . أوجد الوسط التوافقى . يستخدم الوسط التوافقى μ_H أساساً لإيجاد متوسط المعدلات :

$$\mu_H = \frac{N}{\sum(1/X)} \qquad (٢ - ٢٠)$$

$$\mu_H = \frac{2}{(1/60) + (1/15)} = \frac{2}{(1+4)/60} = \frac{2}{5/60} = 2\,\frac{60}{5} = \frac{120}{5} = 24 \text{ mi/h}$$

بينما أن الوسط الحسابى $\mu = \Sigma X/N = (60 + 15)/2 = 75/2 = 37.5$ mi /h لاحظ أنه إذا كان متوسط سرعة المسافرة 37.5 mi / h فإنه يلزمها 60 min (20 m / 37.5 mi) = 32 min لقطع مسافة 20 mi . ولكنها تستغرق 10 min على الطريق خارج المدينة(10 mi بمعدل 60 mi / h) و 40 min على الطريق داخل المدينة(10 mi بمعدل 15 ml / h)فيكون مجموع الزمن 50 min ، والإجابة الصحيحة نحصل عليها باستخدام $\mu_H = 24$ mi / h أى (20 mi / 24 min / h) × 60 min = 50 min

٢ – ١٠ (أ) بالنسبة للبيانات غير المبوبة الواردة فى جدول (٢ – ٧) ، أوجد الربيع الأول والثانى والثالث ، أوجد كذلك المشير الثالث والمئين الستين (ب) أوجد نفس المقاييس للبيانات المبوبة فى جدول (٢ – ١٢) . (وتقسم الربيعات البيانات إلى أربعة أجزاء ، وتقسمها العشيرات إلى عشرة أجزاء ، بينما المئينات تقسمها إلى مائة جزء) .

(أ) 4 = Q_1 (الربيع الأول) (متوسط قيمتى المفردتين العاشرة والحادية عشرة فى جدول ٢–٨)

6 = Q_2 (الربيع الثانى) = قيمة العنصر الذى ترتيبه 20.5 = الوسيط

7.5 = Q_3 (الربيع الثالث) = قيمة العنصر الذى ترتيبه 30.5

5 = D_3 (العشير الثالث) = قيمة العنصر الذى ترتيبه 12.5

7 = P_{60} (المئين الستون) = قيمة العنصر الذى ترتيبه 24.5

(ب) (٢ - ٢١)

$$Q_1 = L + \frac{n/4 - F}{f_1}c$$
$$= \$3.80 + \frac{25/4 - 5}{4}(\$0.10) = \$3.80 + \$0.03125 \cong \$3.83$$

(٢ - ٢٢)

$$Q_2 = L + \frac{n/2 - F}{f_2}c$$
$$= \$3.90 + \frac{25/2 - 9}{5}(\$0.10) = \$3.90 + \$0.07 = \$3.97 = \text{الوسيط}$$

(٢ - ٢٣)

$$Q_3 = L + \frac{3n/4 - F}{f_3}c$$
$$= \$4.00 + \frac{75/4 - 14}{6}(\$0.10) = \$4.00 + \$0.0792 \cong \$4.08$$

(٢ - ٢٤)

$$D_3 = L + \frac{3n/10 - F}{f_3}c$$
$$= \$3.80 + \frac{75/10 - 5}{4}(\$0.10) = \$3.80 + \$0.0625 = \$3.86$$

(٢ - ٢٥)

$$P_{60} = L + \frac{60n/100 - F}{f_{60}}c$$
$$= \$4.00 + \frac{1500/100 - 14}{6}(\$0.10) = \$4.00 + \$0.0167 \cong \$4.02$$

مقاييس التشتت :

٢ - ١١ (أ) أوجد المدى للبيانات غير المبوبة فى جدول (٢ - ٧) .

(ب) أوجد المدى للبيانات غير المبوبة فى جدول (٢ - ١٠) وللبيانات المبوبة فى جدول ٢ - ١٢ .

(ج) ماهى مزايا وعيوب المدى .

(أ) المدى لبيانات غير مبوبة يساوى أكبر قيمة للمشاهدات مطروحاً منها أصغر قيمة للمشاهدات فى مجموعة البيانات . مدى البيانات غير المبوبة فى جدول (٢ - ٧) هو من 2 إلى 10 أى 8 درجات .

(ب) المدى للبيانات غير المبوبة فى جدول (٢ - ١٠) هو من 3.55 $ إلى 4.26 $ أى 0.71 $. ويمتد المدى للبيانات المبوبة من الحد الأدنى للفئة الصغرى إلى الحد الأعلى للفئة الكبرى . المدى للبيانات المبوبة فى جدول (٢ - ١٢) يمتد من 3.50 $ إلى 4.29 $.

(ج) مزايا المدى أنه سهل الحساب والفهم . أما عيوبه فهى أنه يأخذ فى الاعتبار فقط أصغر وأكبر قيمة للتوزيع ، ومن ثم فإنه يتأثر كثيراً بالقيم المتطرفة ، كذلك لايمكن حسابه للتوزيعات المفتوحة . هذه العيوب للمدى تجعل فائدته محدودة (باستثناء استخدامه فى مراقبة الجودة) .

٢ - ١٢ أوجد المدى الربيعى والانحراف الربيعى (نصف المدى الربيعى)

(أ) للبيانات غير المبوبة فى جدول (٢ - ٧) ، (ب) للبيانات المبوبة فى جدول (٢ - ١٢) .

(أ) المدى الربيعى هو الفرق بين الربيع الثالث والربيع الأول أى

$$IR = Q_3 - Q_1 \qquad (٢ - ٢٦)$$

المدى الربيعى للبيانات غير المبوبة فى جدول (٢ - ٧) ، $IR = 7.5 - 4 = 3.5$ (باستخدام قيم Q_3 و Q_1 السابق إيجادها فى مسألة (٢ - ١٠ (أ)) لاحظ أن المدى الربيعى لايتأثر بالقيم المتطرفة لأنها تستخدم فقط النصف

الأوسط البيانات . ومن هنا فإنه أفضل من المدى ، ولكن استخدامه ليس بدرجة شيوع استخدام المقاييس الأخرى للتشتت . أما الانحراف الربيعى .

$$QD = \frac{Q_3 - Q_1}{2} \qquad (٢ - ٢٧)$$

ومن ثم فإن الانحراف الربيعى $QD = (7.5 - 4)/2 = 3.5/2 = 1.75$ ويقيس الانحراف الربيعى المدى المتوسط لربع البيانات .

(ب) المدى الربيعى $IR = Q_3 - Q_1 = \$4.08 - \$3.83 = \$0.25$ (باستخدام قيم Q_3 و Q_1 السابق إيجادها فى مسألة (٢ – ١٠) (ب)) :

$$QD = \frac{Q_3 - Q_1}{2} = \frac{\$4.08 - \$3.83}{2} = \$0.125$$

٢ – ١٣ أوجد الانحراف المتوسط (أ) للبيانات غير المبوبة فى جدول (٢ – ٧) . (ب) للبيانات المبوبة فى جدول (٢ – ٩) .

(أ) حيث $\mu = 6$ (انظر مسألة ٢ – ٣ (أ)) ، فإن

$$\sum|X-\mu| = 1+1+0+4+2+1+0+1+3+3+4+2+1+1+2+0+1+2+2+4$$
$$+3+1+0+1+3+2+4+2+1+3+2+0+1+2+3+0+1+3+4+1$$
$$= 72$$

$$AD = \frac{\sum|X-\mu|}{N} = \frac{72}{40} = 1.8 \text{ درجة}$$

لاحظ أن الانحراف المتوسط يأخذ جميع المشاهدات فى الاعتبار . أى أنه يقيس متوسط الانحراف المطلق لكل مشاهدة عن الوسط الحسابى . ونأخذ القيمة المطلقة (المشار إليها باستخدام خطين رأسين)

إذ أن $\Sigma(X - \mu) = 0$ (أنظر مثال ٥)

(ب) يمكن إيجاد الانحراف المتوسط لنفس البيانات المبوبة بالاستعانة بجدول (٢ – ١٦) :

$$AD = \frac{\sum F|X-\mu|}{N} = \frac{72}{40} = 1.8 \text{ درجات}$$

ويعطى نفس القيمة السابق الحصول عليها من البيانات غير المبوبة .

جدول (٢ – ١٦) حسابات الانحراف المتوسط للبيانات المبوبة فى جدول (٢ – ٩)

الدرجة	مركز الفئة X	التكرار f	الوسط الحسابى μ	$X-\mu$	$\|X-\mu\|$	$f\|X-\mu\|$
1.5–2.4	2	3	6	−4	4	12
2.5–3.4	3	3	6	−3	3	9
3.5–4.4	4	5	6	−2	2	10
4.5–5.4	5	5	6	−1	1	5
5.5–6.4	6	6	6	0	0	0
6.5–7.4	7	8	6	1	1	8
7.5–8.4	8	4	6	2	2	8
8.5–9.4	9	4	6	3	3	12
9.5–10.4	10	2	6	4	4	8
		$\sum f = N = 40$				$\sum\|X-\mu\| = 72$

٢ – ١٤ أوجد الانحراف المتوسط للبيانات المبوبة فى جدول ٢ – ١٢ .

يمكن إيجاد متوسط الانحراف للبيانات المبوبة عن الأجر بالساعة الواردة فى جدول (٢ – ١٢) بالاستعانة بجدول (٢ – ١٧) [$\bar{X} = \$3.95$] (انظر مسألة ٢ – ٤ (ب)) :

$$AD = \frac{\sum f|X - \bar{X}|}{n} = \frac{\$3.60}{25} = \$0.144$$

لاحظ أن الانحراف المتوسط المحسوب للبيانات المبوبة هو تقدير للانحراف المتوسط « الحقيقى » والذى يمكن إيجاده من البيانات غير المبوبة . وعادة ما يختلف قليلا عن الانحراف المتوسط الحقيقى لأننا نستخدم تقدير الوسط الحسابى من البيانات المبوبة فى حساباتنا (قارن قيمتى $\bar{X}$ السابق إيجادهما فى مسألتى (٢ – ٤) (أ) ، (ب)) .

جدول (٢ – ١٧) حسابات الانحراف المتوسط للبيانات المبوبة فى جدول (٢ – ١٢)

أجر الساعة	مركز الفئة X	التكرار f	الوسط الحسابى $\bar{X}$	$X - \bar{X}$	$\|X - \bar{X}\|$	$f\|X - \bar{X}\|$
$3.50–3.59	$3.55	1	$3.95	− $0.40	$0.40	$0.40
3.60–3.69	3.65	2	3.95	−0.30	0.30	0.60
3.70–3.79	3.75	2	3.95	−0.20	0.20	0.40
3.80–3.89	3.85	4	3.95	−0.10	0.10	0.40
3.90–3.99	3.95	5	3.95	0.00	0.00	0.00
4.00–4.09	4.05	6	3.95	0.10	0.10	0.60
4.10–4.19	4.15	3	3.95	0.20	0.20	0.60
4.20–4.29	4.25	2	3.95	0.30	0.30	0.60
		$\sum f = n = 25$				$\sum f\|X - \bar{X}\| = \$3.60$

٢ – ١٥ أوجد التباين والانحراف المعيارى (أ) للبيانات المبوبة فى جدول (٢ – ٧) (ب) للبيانات المبوبة فى جدول (٢ – ٩) (ج) بماذا يمتاز الانحراف المعيارى عن التباين ؟

(a) $\sigma^2 = \frac{\sum (X - \mu)^2}{N}$ و $\mu = 6$ (أنظر مسألة) ٢ – ٣

$$\sum (X - \mu)^2 = 1 + 1 + 0 + 16 + 4 + 1 + 0 + 1 + 9 + 9 + 16 + 4 + 1 + 1 + 4 + 0 + 1 + 4 + 4 + 16 + 9 + 1 + 0 + 1 + 9 + 4 + 16 + 4 + 1 + 9 + 4 + 0 + 1 + 4 + 9 + 0 + 1 + 9 + 16 + 1 = 192$$

$$\sigma^2 = \frac{\sum (X - \mu)^2}{N} = \frac{192}{40} = 4.8$$ درجات مربعة

$$\sigma = \sqrt{\frac{\sum (X - \mu)^2}{N}} = \sqrt{\frac{192}{40}} = \sqrt{4.8} \cong 2.19$$ درجات

(ب) يمكن إيجاد التباين والانحراف المعيارى للبيانات المبوبة للدرجات بالاستعانة بجدول (٢ – ١٨) :

$$\sigma^2 = \frac{\sum f(X - \mu)^2}{N} = \frac{192}{40} = 4.8$$ درجات مربعة

$$\sigma = \sqrt{\sigma^2} = \sqrt{4.8} \cong 2.19$$ درجات

وهى نفس القيم السابق إيجادها من البيانات غير المبوبة

جدول (٢ – ١٨) حسابات التباين والانحراف المعيارى للبيانات فى جدول (٢ – ٩)

الدرجة	مركز الفئة X	التكرار f	الوسط الحسابى μ	$X-\mu$	$(X-\mu)^2$	$f(X-\mu)^2$
1.5–2.4	2	3	6	−4	16	48
2.5–3.4	3	3	6	−3	9	27
3.5–4.4	4	5	6	−2	4	20
4.5–5.4	5	5	6	−1	1	5
5.5–6.4	6	6	6	0	0	0
6.5–7.4	7	8	6	1	1	8
7.5–8.4	8	4	6	2	4	16
8.5–9.4	9	4	6	3	9	36
9.5–10.4	10	2	6	4	16	32
		$\sum f = N = 40$				$\sum f(X-\mu)^2 = 192$

(جـ) يمتاز الانحراف المعيارى عن التباين بأنه يعبر عنه باستخدام نفس وحدات القياس كما فى البيانات بينما يكون تمييز التباين « وحدات القياس مربعة » ويعتبر الانحراف المعيارى أكثر مقاييس التشتت (المطلق) شيوعاً .

٢ – ١٦ أوجد التباين والانحراف المعيارى للبيانات المبوبة فى جدول (٢ – ١٠) .

يمكن إيجاد التباين والانحراف المعيارى للبيانات المبوبة لأجر الساعة بالاستعانة بجدول (٢ – ١٩) ($\bar{X} = \$3.95$ انظر مسألة (٢–٤) (ب)) :

$$s^2 = \frac{\sum f(X-\bar{X})^2}{n-1} = \frac{0.82}{24} \cong 0.0342 \quad \text{دولارات مربعة}$$

$$\text{و} \quad s = \sqrt{\frac{\sum f(X-\bar{X})^2}{n-1}} = \sqrt{0.0342} = \$0.18$$

جدول (٢ – ١٩) حسابات التباين والانحراف المعيارى لبيانات جدول (٢ – ١٢)

أجر الساعة	مركز الفئة X	التكرار f	الوسط الحسابى $\bar{X}$	$X-\bar{X}$	$(X-\bar{X})^2$	$f(X-\bar{X})^2$
\$3.50–3.59	\$3.55	1	\$3.95	−\$0.40	0.16	0.16
3.60–3.69	3.65	2	3.95	−0.30	0.09	0.18
3.70–3.79	3.75	2	3.95	−0.20	0.04	0.08
3.80–3.89	3.85	4	3.95	−0.10	0.01	0.04
3.90–3.99	3.95	5	3.95	0.00	0.00	0.00
4.00–4.09	4.05	6	3.95	0.10	0.01	0.06
4.10–4.19	4.15	3	3.95	0.20	0.04	0.12
4.20–4.29	4.25	2	3.95	0.30	0.09	0.18
		$\sum f = n = 25$				$\sum f(X-\bar{X})^2 = 0.82$

لاحظ أنه فى معادلتى حساب s^2 و s ، يتم استخدام $n-1$ فى المقام بدلا من n . ويرجع ذلك إلى أنه إذا أخذنا عدداً كبيراً من العينات من المجتمع ، فإن متوسط تباين العينات لن يتجه لأن يكون مساوياً لتباين المجتمع σ^2 إلا إذا استخدمنا $n-1$ فى المقام عند حساب s^2 (سوف نناقش ذلك ثانية فى الفصل الخامس) . هذا بالإضافة إلى أن s^2 و s للبيانات المبوبة هى تقديرات للقيم الحقيقية s^2 و s التى يمكن إيجادها من البيانات غير المبوبة لأننا نستخدم فى حساباتنا تقدير $\bar{X}$ من البيانات المبوبة .

٢ — ١٧ بدءاً بمعادلات σ^2 و s^2 المعطاة فى قسم (٢ — ٣) أثبت أن

(أ) (٢ — ٢٨ أ ، ب)

$$\sigma^2 = \frac{\sum X^2 - N\mu^2}{N} \quad \text{و} \quad s^2 = \frac{\sum X^2 - n\bar{X}^2}{n-1}$$

(ب) (٢ — ٢٩ أ ، ب)

$$\sigma^2 = \frac{\sum fX^2 - N\mu^2}{N} \quad \text{و} \quad s^2 \cong \frac{\sum fX^2 - n\bar{X}^2}{n-1}$$

(أ)

$$\sigma^2 = \frac{\sum (X-\mu)^2}{N} = \frac{\sum (X^2 - 2X\mu + \mu^2)}{N} = \frac{\sum X^2 - 2\mu \sum X + N\mu^2}{N}$$

$$= \frac{\sum X^2}{N} - 2\mu^2 + \mu^2 = \frac{\sum X^2 - N\mu^2}{N}$$

ويمكن الحصول على s^2 باستبدال μ واستخدام $\bar{X}$ بدلا منها ، وكذلك باستخدام n بدلا من N فى البسط ، واستخدام $n-1$ بدلا من N فى المقام فى معادلة σ^2

(ب)

$$\sigma^2 = \frac{\sum f(X-\mu)^2}{N} = \frac{\sum f(X^2 - 2X\mu + \mu^2)}{N} = \frac{\sum fX^2 - 2\mu \sum fX + N\mu^2}{N}$$

$$= \frac{\sum fX^2}{N} - 2\mu^2 + \mu^2 = \frac{\sum fX^2 - N\mu^2}{N}$$

ويمكن الحصول على s^2 بنفس الطريقة كما فى (أ) . وتستخدم الصيغ السابقة لتبسيط الحسابات اللازمة لإيجاد σ^2 و s^2 لمجموعة كبيرة من البيانات . ويمكن التبسيط أيضاً باستخدام الترميز (انظر مسألة (٢ — ٦)) .

٢ — ١٨ أوجد التباين والانحراف المعيارى (أ) للبيانات غير المبوبة فى جدول (٢ — ٧ و(ب) للبيانات المبوبة فى جدول (٢ — ٩) باستخدام المعادلات المبسطة مسألة (٢ — ١٧) .

(أ)

$$\sum X^2 = 49 + 25 + 36 + 4 + 64 + 49 + 36 + 49 + 9 + 81 + 100 + 16 + 25 + 25$$
$$+ 16 + 36 + 49 + 18 + 64 + 4 + 9 + 25 + 36 + 49 + 81 + 64 + 4 + 16 + 49$$
$$+ 81 + 16 + 36 + 49 + 64 + 9 + 36 + 49 + 81 + 100 + 25$$
$$= 1{,}632$$

$$\mu = \frac{\sum X}{N} = \frac{240}{40} = 6$$

$$\sigma^2 = \frac{\sum X^2 - N\mu^2}{N} = \frac{1{,}632 - (40)(36)}{40} = \frac{1{,}632 - 1{,}440}{40} = \frac{192}{40} = 4.8$$ درجات مربعة

$$\sigma = \sqrt{\sigma^2} = \sqrt{4.8} \cong 2.19$$ درجات

وهى نفس القيم السابق إيجادها فى مسألة (٢ — ١٥) (أ)

(ب) يمكن إيجاد σ^2 و σ للبيانات المبوبة فى جدول (٢ – ٩) بالاستعانة بجدول (٢ – ٢٠) :

$$\mu = \frac{\sum fX}{N} = \frac{240}{6} = 6$$

$$\sigma^2 = \frac{\sum fX^2 - N\mu^2}{N} = \frac{1{,}632 - (40)(36)}{40} = \frac{1{,}632 - 1{,}440}{40} = \frac{192}{40} = 4.8 \quad \text{درجات مربعة}$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{4.8} \cong 2.19 \quad \text{درجات}$$

وهى نفس القيم السابق الحصول عليها فى (أ) وفى مسألة (٢ – ١٥) .

جدول (٢ – ٢٠) حسابات التباين والانحراف المعيارى للبيانات المبوبة فى جدول (٢ – ٩)

الدرجة	مركز الفئة X	التكرار f	fX	X^2	fX^2
1.5–2.4	2	3	6	4	12
2.5–3.4	3	3	9	9	27
3.5–4.4	4	5	20	16	80
4.5–5.4	5	5	25	25	125
5.5–6.4	6	6	36	36	216
6.5–7.4	7	8	56	49	392
7.5–8.4	8	4	32	64	256
8.5–9.4	9	4	36	81	324
9.5–10.4	10	2	20	100	200
		$\sum f = N = 40$	$\sum fX = 240$		$\sum fX^2 = 1{,}632$

٢ – ١٩ أوجد التباين والانحراف المعيارى للبيانات المبوبة فى جدول (٢ – ١٢) باستخدام المعادلات المبسطة الواردة فى مسألة (٢ – ١٧) (ب)

يمكن إيجاد s^2 و s للبيانات المبوبة فى جدول (٢ – ١٢) بالاستعانة بجدول (٢ – ٢١) :

$$\bar{X} = \frac{\sum fX}{n} = \frac{98.75}{25} = \$3.95$$

$$s^2 = \frac{\sum fX^2 - n\bar{X}^2}{n-1} = \frac{390.8825 - (25)(15.6025)}{24} = \frac{390.8825 - 390.0625}{24} = \frac{0.82}{24}$$

$$\cong 0.0342 \quad \text{دولارات مربعة}$$

و $s \cong \sqrt{0.0342} \cong \$0.18$

وهى نفس القيم السابق الحصول عليها فى مسألة (٢ – ١٦)

جدول ٢ – ٢١ حسابات التباين والانحراف المعيارى للبيانات المبوبة فى جدول (٢ – ١٢)

أجر الساعة	مركز الفئة X	التكرار f	fX	X^2	fX^2
\$3.50–3.59	\$3.55	1	\$ 3.55	12.6025	12.6025
3.60–3.69	3.65	2	7.30	13.3225	26.6450
3.70–3.79	3.75	2	7.50	14.0625	28.1250
3.80–3.89	3.85	4	15.40	14.8225	59.2900
3.90–3.99	3.95	5	19.75	15.6025	78.0125
4.00–4.09	4.05	6	24.30	16.4025	98.4150
4.10–4.19	4.15	3	12.45	17.2225	51.6675
4.20–4.29	4.25	2	8.50	18.0625	36.1250
		$\sum f = n = 25$	$\sum fX = \$98.75$		$\sum fX^2 = 390.8825$

٢ – ٢٠ أوجد معامل الاختلاف V (أ) لبيانات جدول (٢ – ٧) (ب) لبيانات جدول (٢ – ١٢) (جـ) مافائدة معامل الاختلاف ؟

(أ) عندما $\mu = 6$ $\sigma \cong 2.19$ (انظر مسألة (٢ – ١٨))

$$V = \frac{\sigma}{\mu} \cong \frac{2.19 \text{ points}}{6 \text{ points}} \cong 0.635, \quad \text{or} \quad 6.35\%$$

(ب) عندما $\bar{X} = \$3.95$ $s \cong \$0.18$ (انظر مسألة (٢ – ١٩))

$$V = \frac{s}{\bar{X}} \cong \frac{\$0.18}{\$3.95} \cong 0.046, \quad \text{or} \quad 4.6\%$$

(جـ) يقيس معامل الاختلاف التشتت النسبى فى البيانات ويعبر عنه بأرقام مطلقة وليس بوحدات . وذلك على عكس الانحراف المعيارى والمقاييس الأخرى للتشتت المطلق والتى تأخذ أرقامها المحسوبة نفس وحدات البيانات الأصلية . ومن ثم فان معامل الاختلاف يمكن استخدامه لمقارنة تشتت توزيعين أو أكثر إذا ما اختلفت وحدات القياس أو اختلفت متوسطاتها الحقيقية. فعلى سبيل المثال يمكننا القول أن تشتت بيانات جدول (٢ – ٧) أكبر من تشتت بيانات جدول (٢–١٢) . كما يمكن استخدام معامل الاختلاف لمقارنة تشتت نفس نوع البيانات على مدى عدة فترات زمنية (عندما تتغير μ أو $\bar{X}$ وتتغير σ أو s) .

شكل التوزيعـــات التكـــرارية :

٢ – ٢١ أوجد معامل بيرسون للالتواء للبيانات (المبوبة) (أ) فى جدول (٢ – ٩) و (ب) فى جدول (٢ – ١٢)

(أ) عندما $\mu = 6$ ، الوسيط $= 6.17$ (انظر مسألة (٢ – ٣) (ب)) ، $\sigma \cong 2.19$ (انظر مسألة (٢ – ١٥ (ب)) .

$$Sk = \frac{3(\mu - \text{med})}{\sigma} \cong \frac{3(6 - 6.17)}{2.19} \cong \frac{3(-0.17)}{2.19} \cong -0.23$$ (عدد مطلق)

لاحظ أن الوسيط أكبر من الوسط الحسابى وأن التوزيع سالب الالتواء قليلا (انظر شكل (٢ – ٥ جـ)).

(ب) عندما $\bar{X} = \$\ 3.95$ ، الوسيط $\$\ 3.97$ (انظر مسألة (٢ – ٤ (ب)) ، $s \cong \$0.18$ (انظر مسألة (٢ – ١٦))

$$Sk = \frac{3(\bar{X} - \text{med})}{s} \cong \frac{3(3.95 - 3.97)}{0.18} = \frac{3(-0.02)}{0.18} = -0.33$$

(انظر شكل ٢ – ٦ جـ)

٢ – ٢٢ باستخدام صيغة الالتواء المؤسسة على العزم الثالث أوجد معامل الالتواء لبيانات

(أ) جدول (٢ – ٩) (ب) جدول (٢ – ١٢)

(أ) يمكن إيجاد معامل الالتواء لبيانات جدول (٢ – ٩) باستخدام الصيغة المبينة على العزم الثالث بالاستعانة بجدول (٢ – ٢٢) :

$$SK = \frac{\Sigma f(X-\mu)^3}{N\sigma^3} \cong \frac{-1.05}{2.19^3} = -0.1$$

ويشير هذا إلى أن التوزيع سالب الالتواء ولكن درجة الالتواء تقاس بأسلوب مختلف عنه فى مسألة (٢ – ٢١)

جدول (٢ – ٢٢) حسابات الالتواء لبيانات جدول (٢ – ٩) .

الدرجة	مركز الفئة X	التكرار f	الوسط الحسابى μ	$X-\mu$	$(X-\mu)^3$	$f(X-\mu)^3$
1.5–2.4	2	3	6	−4	−64	−192
2.5–3.4	3	3	6	−3	−27	−81
3.5–4.4	4	5	6	−2	−8	−40
4.5–5.4	5	5	6	−1	−1	−5
5.5–6.4	6	6	6	0	0	0
6.5–7.4	7	8	6	1	1	8
7.5–8.4	8	4	6	2	8	32
8.5–9.4	9	4	6	3	27	108
9.5–10.4	10	2	6	4	64	128
		$\sum f = N = 40$				$\sum f(X-\mu)^3 = -42$

(ب) انظر جدول (٢ – ٢٣)

$$SK = \frac{\Sigma f(X-\bar{X})^3}{n s^3} = \frac{-0.0216}{0.18^3} = -0.36$$

لاحظ أنه بصرف النظر عن مقياس الالتواء المستخدم ، فإن توزيعات بيانات جدول (٢ – ٩) وجدول (٢ – ١٢) سالبة الالتواء ، والتواء الأخير أكبر من الأول .

جدول (٢ – ٢٣) حسابات الالتواء لبيانات جدول (٢ – ١٢)

أجر الساعة	مركز الفئة X	التكرار f	الوسط الحسابى $\bar{X}$	$X-\bar{X}$	$(X-\bar{X})^3$	$f(X-\bar{X})^3$
$3.50–3.59	$3.55	1	$3.95	−$0.40	−0.064	−0.064
3.60–3.69	3.65	2	3.95	−0.30	−0.027	−0.054
3.70–3.79	3.75	2	3.95	−0.20	−0.008	−0.016
3.80–3.89	3.85	4	3.95	−0.10	−0.001	−0.004
3.90–3.99	3.95	5	3.95	0	0	0
4.00–4.09	4.05	6	3.95	0.10	0.001	0.006
4.10–4.19	4.15	3	3.95	0.20	0.008	0.024
4.20–4.29	4.25	2	3.95	0.30	0.027	0.054
						$\sum f(X-\bar{X})^3 = -0.054$

٢ – ٢٣ أوجد معامل التفرطح لبيانات (أ) جدول (٢ – ٩) (ب) جدول (٢ – ١٢)

(أ) يمكن إيجاد معامل التفرطح لبيانات جدول (٢ – ٩) بالاستعانة بجدول ٢ – ٢٤ :

$$\text{معامل التفرطح} = \frac{\Sigma f(X-\mu)^4}{N\sigma^4} \quad \frac{50.1}{2.19^4} \cong 2.18 \quad (\text{عدد مطلق})$$

أى أن التوزيع مفرطح (انظر شكل ٢ – ٥ جـ) .

جدول (٢ – ٢٤) حسابات التفرطح لبيانات جدول (٢ – ٩)

$f(X-\mu)^4$	$(X-\mu)^4$	$X-\mu$	الوسط الحسابى μ	التكرار f	مركز الفئة X	الدرجة
768	256	-4	6	3	2	1.5–2.4
243	81	-3	6	3	3	2.5–3.4
80	16	-2	6	5	4	3.5–4.4
5	1	-1	6	5	5	4.5–5.4
0	0	0	6	6	6	5.5–6.4
8	1	1	6	8	7	6.5–7.4
64	16	2	6	4	8	7.5–8.4
324	81	3	6	4	9	8.5–9.4
512	256	4	6	2	10	9.5–10.4
$\sum f(X-\mu)^4 = 2{,}004$				$\sum f = N = 40$		

(ب) بالاستعانة بجدول (٢ – ٢٥) :

$$\text{معامل التفرطح} = \frac{\Sigma\, f\,(X - X)^4}{n\, s^4} \cong \frac{0.00268}{0.001} \cong 2.68$$

أى أن توزيع الأجور أيضاً مفرطح (انظر شكل ٢ – ٦ جـ)

جدول (٢ – ٢٥) حسابات معامل التفرطح لبيانات جدول (٢ – ١٢)

$f(X-\bar{X})^3$	$(X-\bar{X})^3$	$X-\bar{X}$	الوسط الحسابى $\bar{X}$	التكرار f	مركز الفئة X	أجر الساعة
-0.064	-0.064	$-\$0.40$	\$3.95	1	\$3.55	\$3.50–3.59
-0.054	-0.027	-0.30	3.95	2	3.65	3.60–3.69
-0.016	-0.008	-0.20	3.95	2	3.75	3.70–3.79
-0.004	-0.001	-0.10	3.95	4	3.85	3.80–3.89
0	0	0	3.95	5	3.95	3.90–3.99
0.006	0.001	0.10	3.95	6	4.05	4.00–4.09
0.024	0.008	0.20	3.95	3	4.15	4.10–4.19
0.054	0.027	0.30	3.95	2	4.25	4.20–4.29
$\sum f(X-\bar{X})^3 = -0.054$						

مسائل اضافيه

التـــوزيعـــات التكـــرارية :

٢ – ٢٤ يوضح جدول (٢ – ٢٦) التوزيع التكرارى لأسعار البنزين فى 48 محطة بإحدى المدن . اعرض البيانات فى شكل مدرج تكرارى ، مدرج تكرارى نسبى ، مضلع تكرارى ، ومنحنى متجمع .

جدول (٢ – ٢٦) التوزيع التكرارى لأسعار البنزين

التكرار	السعر
4	\$1.00–1.04
6	1.05–1.09
10	1.10–1.14
15	1.15–1.19
8	1.20–1.24
5	1.25–1.29

٢ – ٢٥ جدول (٢ – ٢٧) يوضح التوزيع التكرارى لدخول عينة مكونة من 100 أسرة مأخوذة من إحدى المدن . اعرض البيانات باستخدام مدرج تكرارى ، مدرج تكرارى نسبى ، مضلع تكرارى ، ومنحنى متجمع

جدول (٢ – ٢٧) التوزيع التكرارى لدخل العائلة

التكرار	دخل العائلة
$10,000–11,999	12
12,000–13,999	14
14,000–15,999	24
16,000–17,999	15
18,000–19,999	13
20,000–21,999	7
22,000–23,999	6
24,000–25,999	4
26,000–27,999	3
28,000–29,999	2
	100

مقاييس النزعة المركزية :

٢ – ٢٦ أوجد (أ) الوسط الحسابى (ب) الوسيط (جـ) المنوال للبيانات المبوبة فى جدول (٢ – ٢٦) .

الإجابة : (أ) الوسط الحسابى = 1.15 $ (ب) الوسيط 1.16 $ (جـ) المنوال = 1.17 $

٢ – ٢٧ أوجد (أ) الوسط الحسابى (ب) الوسيط (جـ) المنوال التوزيع التكرارى للدخل فى جدول (٢ – ٢٧) .

الإجابة : (أ) الوسط الحسابى 17, 000 $ (ب) الوسيط = 16,000 $ (جـ) المنوال = 15, 053 $

٢ – ٢٨ أوجد الوسط الحسابى للبيانات المبوبة فى (أ) جدول (٢ – ٢٦) و(ب) جدول (٢–٢٧) باستخدام الترميز

الإجابة : (أ) $\mu = \$1.15$ (ب) $\bar{X} = \$\,17,000$

٢ – ٢٩ تدفع شركة أجر 5/12 من قوة العمل بها بمعدل 5$ للساعة ، أجر 1/3 قوة العمل بمعدل 6 $ للساعة ، وأجر 1/4 قوة العمل بمعدل S 7 فى الساعة . ماهو المتوسط المرجح للأجور المدفوعة بالشركة ؟

الإجابة : $\mu_w \cong \$5.83$

٢ – ٣٠ حصل مستثمر على عائد من رأسماله المستثمر قدره 1% للسنة الأولى ، 4 % للسنة الثانية ، 16 % للسنة الثالثة .

(أ) أوجد μ_G (ب) أوجد μ (جـ) أى منهما يعتبر مقياساً مناسباً ؟

الإجابة : (أ) $\mu_G = 4\%$ (ب) $\mu = 7\%$ (جـ) μ_G

٢ – ٣١ قطعت طائرة مسافة 200 ميل بمعدل 600 ميل / ساعة ومسافة 100 ميل بمعدل 500 ميل / ساعة . ماهو متوسط السرعة ؟

الإجابة : 562.5 ميل / ساعة .

٢ – ٣٢ يشترى سائق سيارة ماقيمته 10 $ من البنزين بسعر 0.90 $ للجالون ، وما قيمته 10 $ بسعر 1.10 $. ماهو متوسط سعر الجالون .

الإجابة : 0.99 $ للجالون .

٢ – ٣٣ للبيانات المبوبة فى جدول (٢ – ٢٦) أوجد (أ) الربيع الأول (ب) الربيع الثانى (جـ) الربيع الثالث (د) المشير الرابع (هـ) المئين السبعين .

الإجابة : (أ) $Q_1 = \$1.11$ (ب) $Q_2 \cong \$1.16$ (جـ) $Q_3 \cong \$1.21$ (د) $D_4 = \$1.146$
(هـ) $P_{70} \cong \$1.195$

٢ – ٣٤ للبيانات المبوبة فى جدول (٢ – ٢٧) أوجد (أ) الربيع الأول (ب) الربيع الثالث (جـ) المشير الثالث (د) المئين الستين

الإجابة : (أ) $Q_1 \cong \$13,857$ (ب) $Q_3 \cong \$19,538$
(جـ) $D_3 \cong \$14,333$ (د) $P_{60} \cong \$17,333$

مقاييس التشتت :

٢ – ٣٥ ماهو المدى لتوزيع (أ) أسعار البنزين فى جدول (٢ – ٢٦) (ب) دخل الأسرة فى جدول (٢ – ٢٧) ؟

الإجابة : (أ) $0.29 $ (ب) من 10,000 $ إلى 29,999 $ أو 20,000 $

٢ – ٣٦ أوجد المدى الربيعى والانحراف الربيعى لبيانات (أ) جدول (٢ – ٢٦) (ب) جدول (٢ – ٢٧)

الإجابة : (أ) QD ≅ \$0.05 و IR ≅ \$0.10 (ب) QD ≅ \$238 و IR ≅ \$476

٢ – ٣٧ أوجد متوسط الانحراف لبيانات (أ) جدول (٢ – ٢٦) (ب) جدول (٢ – ٢٧)

الإجابة : (أ) 0.0575 $ (ب) 3,520 $

٢ – ٣٨ أوجد (أ) التباين (ب) الانحراف المعيارى للتوزيع التكرارى لأسعار البنزين فى جدول (٢ – ٢٦)

الإجابة : (أ) دولارات مربعة $\sigma^2 \cong 0.0048$ (ب) $\sigma \cong \$0.0693$

٢ – ٣٩ أوجد (أ) التباين (ب) الانحراف المعيارى للتوزيع التكرارى لدخل العائلة فى جدول (٢ – ٢٧)

الإجابة : (أ) دورات مربعة $s^2 = 19,760,000$ (ب) $s \cong \$4,445.22$

٢ – ٤٠ باستخدام الصيغ الحسابية الأسهل أوجد (أ) التباين (ب) الانحراف المعيارى لتوزيع أسعار البنزين فى جدول (٢ – ٢٦) .

الإجابة : (أ) $s^2 = \$\,19,760,000$ (ب) $\sigma \cong \$0.0693$

٢ – ٤١ باستخدام الصيغ الحسابية الأسهل أوجد (أ) التباين (ب) الانحراف المعيارى لدخل الأسرة فى جدول (٢ – ٢٧)

الإجابة : (أ) $s^2 = \$^2\,19,760,000$ (ب) $s \cong \$4,445.22$

٢ – ٤٢ طبقاً لنظرية تشبتشيف ، ماهى أقل نسبة من المشاهدات التى لاتبعد عن الوسط الحسابى بأكثر من : (أ) 1.5 انحراف معيارى (ب) 2.5 انحراف معيارى ؟

الإجابة : (أ) 56% (ب) 84%

٢ – ٤٣ للبيانات غير المبوبة فى جدول (٢ – ٧) (أ) أوجد نسبة المشاهدات التى لاتبعد عن الوسط الحسابى بأكثر من 1.5 انحراف معيارى (ب) هل يتفق هذا ونظرية تشبتشيف ؟ (جـ) أوجد نسبة المشاهدات التى لاتبعد عن الوسط الحسابى بأكثر من 2.5 % انحراف معيارى (د) هل يتفق هذا ونظرية تشبتشيف ؟

الإجابة : (أ) 87.5% (ب) نعم (جـ) 100% (د) نعم

٢ - ٤٤ أوجد معامل الاختلاف (أ) لبيانات جدول (٢ - ٢٦) (ب) لبيانات جدول (٢ - ٢٧) (جـ) أى البيانات أكثر تشتتاً ؟

الإجابة : (أ) 0.060 أو 6% (ب) 0.261 أو 26.1% (جـ) بيانات جدول (٢ - ٢٧)

أشكال التوزيعات التكرارية :

٢ - ٤٥ أوجد معامل بيرسون للالتواء لبيانات (أ) جدول (٢ - ٢٦) (ب) جدول (٢ - ٢٧)

الإجابة : (أ) 0.43 — (ب) 0.67

٢ - ٤٦ أوجد معامل الالتواء باستخدام العزم الثالث لبيانات (أ) جدول (٢ - ٢٦) (ب) جدول (٢ - ٢٧)

الإجابة : (أ) 1.88 — (ب) 755

٢ - ٤٧ أوجد معامل التفرطح لبيانات (أ) جدول (٢ - ٢٦) (ب) جدول (٢ - ٢٧)

الإجابة : (أ) 177 (ب) 300

الفصل الثالث

الاحتمال والتوزيعات الاحتمالية

٣ـ١ احتمال حدث منفرد

إذا كان الحدث A يمكن أن يحدث بطرق عددها n_A من بين نواتج متساوية الفرصة فى الوقوع عددها N ، فإن احتمال وقوع الحدث A يعرف كما يلى :

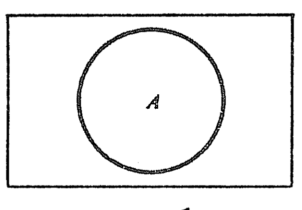
$$P(A) = \frac{n_A}{N} \qquad (٣-١)$$

حيث $P(A)$ = احتمال وقوع الحدث A

A = عدد الطرق التى يمكن أن يقع بها A

N = العدد الكلى للنواتج المتساوية الفرصة فى الوقوع .

ويمكن تصور الاحتمال باستخدام **شكل فن** . فى شكل (٣ – ١) تمثل الدائرة الحدث A ، بينما تمثل المساحة الكلية للمستطيل كل النواتج الممكنة



شكل ٣ – ١

وتتراوح $P(A)$ بين 0 و 1 .

$$0 \leqslant P(A) \leqslant 1 \qquad (٣-٢)$$

فإذا كانت $P(A) = 0$ ، فإن الحدث A لايمكن أن يقع . وإذا كانت $P(A) = 1$ ، فإن الحدث A مؤكد الوقوع .

وإذا استخدمنا $P(A')$ لتمثل احتمال عدم وقوع الحدث A فإن ،

$$P(A) + P(A') = 1 \qquad (٣-٣)$$

مثال (١) : عند إلقاء قطعة نقود متوازنة فإن الصورة والكتابة يمثلان ناتجين لهما نفس فرصة الوقوع ، أى أن

$$P(\mathrm{H}) = \frac{n_\mathrm{H}}{N} = \frac{1}{2}$$

$$P(\mathrm{T}) = \frac{n_\mathrm{T}}{N} = \frac{1}{2}$$

و

$$P(\mathrm{H}) + P(\mathrm{T}) = 1$$

مثال (٢) : عند إلقاء حجر نرد غير متحيز فإن هناك ستة نواتج متساوية الفرصة فى الحدوث 1, 2, 3, 4, 5, 6 ومن ثم فإن ،

$$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = \frac{1}{6}$$

ويكون احتمال عدم ظهور 1 هو

$$P(1') = 1 - P(1) = 1 - \frac{1}{6} = \frac{5}{6}$$

و

$$P(1) + P(1') = \frac{1}{6} + \frac{5}{6} = \frac{6}{6} = 1$$

مثال (٣) : سحبت ورقة واحدة من مجموعة أوراق لعب ذات 52 ورقة ممزوجة مزجاً جيداً . وحيث أن مجموعة أوراق اللعب تحتوى على 4 أولاد فإن احتمال أن تكون الورقة المسحوبة ولداً ، J ،

$$J = \frac{n_J}{N} = \frac{4}{52} = \frac{1}{13}$$

وحيث أن المجموعة تحتوى على 13 ورقة دينارية ، D ، فإن ،

$$P(D') = 1 - P(D) = 1 - \frac{13}{52} = 1 - \frac{1}{4} = \frac{3}{4}$$

و

$$P(D) + P(D') = \frac{1}{4} + \frac{3}{4} = 1$$

مثال (٤) : افترض أننا حصلنا فى مائة رمية لعملة متوازنة على 53 صورة ، 47 كتابة . فإن التكرار النسبى للصورة يكون 100 / 53 أو 0.53 . إن هذا هو **التكرار النسبى أو الاحتمال التجريبى** وينبغى تمييزه عن **الاحتمال الكلاسيكى أو المسبق** $P(H) = 0.5$. ومع تزايد عدد الرميات ليقترب من مالانهاية فإن التكرار النسبي أو الاحتمال التجريبى يقترب من الاحتمال المسبق أو الكلاسيكى . وعلى سبيل المثال فإن التكرار النسبي أو الاحتمال التجريبى قد يكون 0.517 فى حالة 1,000 رمية ، 0.508 فى حالة 10,000 رمية ، وهكذا .

٣ـ٢ احتمال الأحداث المتعددة

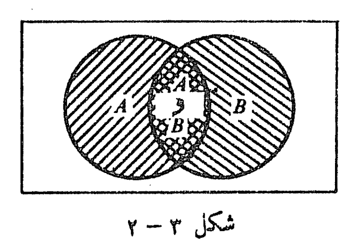
١ – **قاعدة الجمع للأحداث المتنافية** : يعتبر الحدثان A و B **متنافيان بالتبادل** إذا كان وقوع A يحجب وقوع B والعكس بالعكس . عندئذ

$$P(A \text{ أو } B) = P(A) + P(B) \qquad (٣-٤)$$

٢ – **قاعدة الجمع للأحداث غير المتنافية** : يعتبر الحدثان A و B **غير متنافيين** إذا كان وقوع A لايحجب وقوع B والعكس بالعكس فيكون

$$P(A \text{ أو } B) = P(A) + P(B) - P(A \text{ و } B) \qquad (٣-٥)$$

وتطرح قيمة $P(A \text{ و } B)$ حتى نتجنب حسابها مرتين . ويمكن إدراك ذلك من شكل فن الموضح بشكل (٣ – ٢) .



شكل ٣ – ٢

٣ – **قاعدة الضرب للأحداث المستقلة :** يعتبر الحدثان A و B مستقلين إذا كان وقوع A غير مرتبط بأى طريقة بوقوع B . عندئذ الاحتمال المشترك للحدثين A و B هو

$$P(A \text{ و } B) = P(A) \cdot P(B) \tag{٣ – ٦}$$

٤ – **قاعدة الضرب للأحداث غير المستقلة :** يعتبر الحدثان غير مستقلين إذا كان وقوع أحدهما مرتبطاً بطريقة ما بوقوع الآخر . عندئذ

$$P(A \text{ و } B) = P(A) \cdot P(B/A) \tag{٣ – ٧}$$

وتقرأ كالآتى : « احتمال وقوع كل من الحدثين A و B يساوى احتمال وقوع الحدث A مضروباً فى احتمال وقوع الحدث B إذا علم أن الحدث A قد وقع فعلا » .

(٣ – ٨) $P(B/A) =$ الاحتمال الشرطى للحدث B علماً بأن الحدث A قد وقع فعلا .

$$P(A \text{ و } B) = P(B \text{ و } A) \tag{٣ – ٩}$$

انظر مسألة (٣ – ١٥) (جـ) و (د) .

مثال (٥) : فى رمية واحدة لحجر نرد ، يمكن الحصول على واحد من ستة نواتج ممكنة : 1, 2, 3, 3, 4, 5, 6 . هذه الأحداث متنافية بالتبادل . فإذا كانت النردة غير متحيزة فإن $P(1) = P(2) = P(3) = P(4) = P = (5) = P = (6) = 1/6$ ويكون احتمال الحصول على 2 أو 3 فى رمية واحدة للنردة :

$$P(2 \text{ أو } 3) = P(2) + P(3) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$$

وبالمثل :

$$P(2 \text{ أو } 3 \text{ أو } 4) = P(2) + P(3) + P(4) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}$$

مثال (٦) : عند سحب ورقة واحدة عشوائياً من مجموعة أوراق لعب مخلوطة خلطاً جيداً فإن الحدث « بستونى s » والحدث « شايب k » « غير متنافية بالتبادل » لأنه يمكننا سحب شايب بستونى ، فيكون

$$P(\text{S أو K}) = P(\text{S}) + P(\text{K}) - P(\text{S و K}) = \frac{13}{52} + \frac{4}{52} - \frac{1}{52} = \frac{16}{52} = \frac{4}{13}$$

وباستخدام نظرية المجموعات ، فإن العبارة السابقة يمكن إعادة كتابتها فى الصورة

$$P(\text{S} \cup \text{K}) = P(\text{S}) + P(\text{K}) - P(\text{S} \cap \text{K}) = \frac{13}{52} + \frac{4}{52} - \frac{1}{52} = \frac{16}{52} = \frac{4}{13}$$

حيث الرمز $\cup$ (ويقرأ « اتحاد) يحل محل « أو » ، الرمز $\cap$ (ويقرأ « تقاطع ») يحل محل « و » .

مثال (٧) : تعتبر نواتج رميتين متتاليتين لقطعة نقود متوازنة أحداثاً مستقلة . فناتج الرمية الأولى لايؤثر على أى نحو على ناتج الرمية الثانية . فيكون .

$$P(\text{H و H}) = P(\text{H} \cap \text{H}) = P(\text{H}) \cdot P(\text{H}) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \text{ أو } 0.25,$$

وبالمثل

$$P(\text{H و H و H}) = P(\text{H} \cap \text{H} \cap \text{H}) = P(\text{H}) \cdot P(\text{H}) \cdot P(\text{H}) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8} \text{ و } 0.125,$$

مثال (٨) : احتمال الحصول على « شايب دينارى » عند سحب الورقة الأولى من مجموعة أوراق اللعب هو

$$P(\text{K}_\text{D}) = \frac{1}{52}$$

فإذا كانت الورقة الأولى المسحوبة هي فعلا شايب ديناري ، ولم تعد الورقة المسحوبة إلى المجموعة فإن احتمال الحصول على شايب آخر عند سحب الورقة الثانية تتوقف على نتيجة سحب الورقة الأولى ، حيث أصبح بالمجموعة 3 شايب فقط من مجموع 51 ورقة المتبقية فى المجموعة . ويكون الاحتمال الشرطى لسحب شايب آخر ، بمعلومية أن الورقة الأولى كانت شايب ديناري ولم تعد للمجموعة ، هو

$$P(\mathrm{K}/\mathrm{K_D}) = 3/51$$

وإذن فاحتمال سحب شايب ديناري فى السحبة الأولى ، بدون إحلال ، وسحب شايب آخر فى السحبة الثانية هو

$$P(\mathrm{K_D} \text{ و } \mathrm{K}) = P(\mathrm{K_D}) \cdot P(\mathrm{K}/\mathrm{K_D}) = \frac{1}{52} \cdot \frac{3}{51} = \frac{3}{2{,}652}$$

أى حوالى واحد فى الألف . وترتبط بالاحتمال الشرطى نظرية بيز (انظر مسألة ٣ – ١٧) . كما تراجع المسألة ٣ – ١٨ التباديل والتوافيق ، أو « أساليب العد » .

٣–٣ التوزيعات الاحتمالية المنفصلة : توزيع ذى الحدين

المتغير العشوائى : هو متغير ترتبط قيمه باحتمال تحقق تلك القيم . المتغير العشوائى المنفصل (مقارنة بالمتصل) هو المتغير الذى يمكن أن يأخذ فقط قيما محدودة ومتميزة . وتسمى مجموعة كل القيم الممكنة للمتغير العشوائى واحتمالاتها المناظرة **التوزيع الاحتمالى** . ويكون مجموع الاحتمالات 1 (انظر مثال ٩) .

إن التوزيع ذا الحدين هو أحد التوزيعات الاحتمالية المنفصلة . ويستخدم لإيجاد احتمال وقوع حدث معين (نجاح) عدداً من المرات مقداره X من بين n من المحاولات لنفس التجربة (ونرمز لهذا الاحتمال بالرمز $P(X)$) عندما تتحقق الشروط التالية :

(١) هناك فقط ناتجان ممكنان ومتنافيان لكل محاولة (٢) المحاولات وعددها n مستقلة عن بعضها البعض (٣) احتمال وقوع الحدث المعين فى كل محاولة (النجاح) P ، ثابت ولا يتغير من محاولة لأخرى . فيكون

$$P(X) = \frac{n!}{X!(n-X)!} p^X (1-p)^{n-X} \qquad (٣-١٠)$$

حيث $n!$ (وتقرأ « مضروب n ») $= n \cdot (n-1) \cdot (n-2) \ldots 3.2.1$ و $0! = 1$ بالتعريف (انظر مسألة (٣ – ١٨)) .

ويكون متوسط توزيع ذى الحدين

$$\mu = np \qquad (٣-١١)$$

وانحرافه المعيارى

$$\sigma = \sqrt{np(1-p)} \qquad (٣-١٢)$$

فإذا كانت $p = 1 - p = 0.5$ ، فإن توزيع ذى الحدين يكون متماثلا ؛ وإذا كانت $p < 0.5$ ، يكون التوزيع ملتوياً إلى اليمين ؛ وإذا كانت $p > 0.5$ يكون التوزيع ملتوياً إلى اليسار .

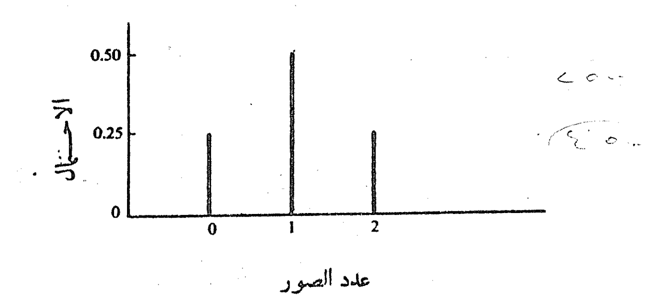
مثال (٩) عند رمى عملة متوازنة مرتين فإن النواتج الممكنة هى TT, TH, HT, HH وإذن

$$P(0\mathrm{H}) = \frac{1}{4} \qquad P(1\mathrm{H}) = \frac{1}{2} \qquad \text{و} \qquad P(2\mathrm{H}) = \frac{1}{4}$$

وهكذا فإن عدد الصور متغير عشوائى منفصل ، وتمثل مجموعة كل النواتج الممكنة مع احتمالاتها المناظرة توزيعاً احتمالياً منفصلا . (انظر جدول ٣ – ١ ، وشكل ٣ – ٣) .

جدول ٣ — ١ التوزيع لعدد الصور فى رميتين لعملة متوازنة

عدد الصور	النواتج الممكنة	احتمال عدد الصور
0	TT	0.25
1	TH, HT	0.50
2	HH	0.25
		1.00



عدد الصور

شكل ٣ — ٣ التوزيع الاحتمالى لعدد الصور فى رميتين لعملة متوازنة

مثال (١٠) : باستخدام توزيع ذى الحدين يمكننا إيجاد احتمال الحصول على 4 صور فى 6 رميات لعملة متوازنة كالآتى :

$$P(4) = \frac{6!}{4!\,(6-4)!}(1/2)^4(1/2)^2 = \frac{6\cdot5\cdot4\cdot3\cdot2\cdot1}{4\cdot3\cdot2\cdot1\cdot2\cdot1}(1/16)(1/4) = 15(1/64) = \frac{15}{64} \cong 0.23$$

ويمكن تجنب الحسابات الطويلة لإيجاد الاحتمالات ، عندما تكون n و X أعداداً كبيرة باستخدام الملحق ١ . إن عدد الصور المتوقع فى ست رميات هو $\mu = np = (6)(1/2) = 3$. ويكون الانحراف المعيارى للتوزيع الاحتمالى لست رميات هو

$$\sigma = \sqrt{np(1-p)} = \sqrt{(6)(1/2)(1/2)} = \sqrt{6/4} = \sqrt{1.5} \cong 1.22 \quad \text{صورة}$$

وهذا التوزيع متماثل لأن $p = 0.5$. وإذا لم تكن التجربة رمياً لعملة ولم تكن المحاولات المتتابعة مستقلة (كما فى حالة المعاينة بدون إحلال) ، كان علينا أن نستخدم **التوزيع فوق الهندسى** (انظر مسألة ٣ — ٢٧) .

٣—٤ توزيع بواسون

توزيع بواسون هو توزيع احتمالى منفصل آخر . ويستخدم لتحديد احتمال وقوع عدد معين من النجاحات فى وحده الزمن ، وذلك عندما تكون الأحداث أو « النجاحات » مستقلة عن بعضها البعض وعندما يبقى متوسط عدد النجاحات ثابتاً لوحدة الزمن . عندئذ

$$P(X) = \frac{\lambda^X e^{-\lambda}}{X!} \qquad (٣ - ١٢)$$

حيث :
X = العدد المعين من النجاحات
$P(X)$ = احتمال عدد X من النجاحات
λ = (الحرف اليونانى لامدا) = متوسط عدد النجاحات فى وحدة الزمن
e = أساس نظام اللوغاريتمات الطبيعى ، أو 2.71828

وبمعلومية قيمة λ (القيمة المتوقعة أو متوسط وتباين توزيع بواسون) ، يمكن إيجاد $e^{-\lambda}$ من ملحق٢ ، والتعويض فى معادلة (٣ — ١٢) ، وإيجاد $P(X)$

مثال (١١) : يتلقى قسم بوليس فى المتوسط 5 مكالمات فى الساعة فيكون احتمال تلقى مكالمتين فى ساعة مختارة عشوائياً هو

$$P(X) = \frac{\lambda^X e^{-\lambda}}{X!} = \frac{5^2 e^{-5}}{2!} = \frac{(25)(0.00674)}{2} = 0.08425$$

ويمكن استخدام توزيع بواسون كتقريب للتوزيع ذى الحدين عندما تكون n كبيرة وتكون P أو $1-p$ صغيرة [$n \geq 30,\ np < 5$ أو $n(1-p) < 5$] انظر المسألة (٣ - ٣٠)

٣-٥ التوزيعات الاحتمالية المتصلة : التوزيع الطبيعى

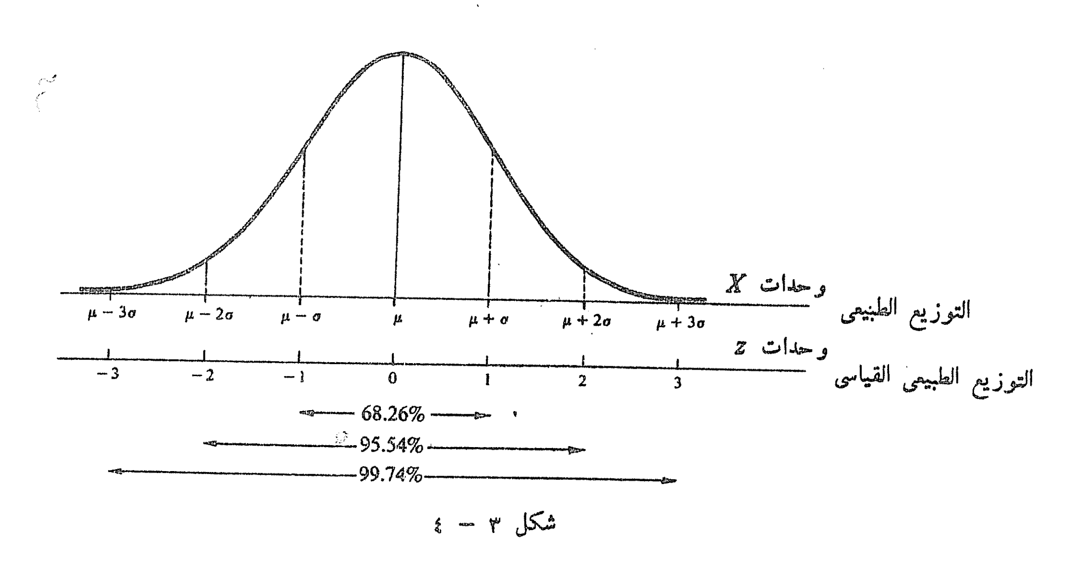
المتغير العشوائى المتصل X هو المتغير الذى يمكن أن يأخذ عدداً لانهائياً من القيم داخل أى فترة معلومة . احتمال أن تقع X داخل أى فترة يمثله مساحة التوزيع الاحتمالى (ويسمى أيضاً دالة الكثافة) داخل هذه الفترة . والمساحة الكلية تحت المنحنى (الاحتمال) تساوى 1 (انظر المسألة ٣ - ٣١) .

التوزيع الطبيعى هو توزيع احتمال متصل وهو أكثر التوزيعات استخداماً فى التحليل الإحصائى (انظر المسألة ٣ - ٣٢) . والتوزيع الطبيعى جرسى الشكل ومتماثل حول الوسط الحسابى . ويمتد إلى مالا نهاية فى الاتجاهين ، ولكن معظم المساحة (الاحتمال) يتركز حول الوسط الحسابى (انظر شكل ٣ - ٤) .

التوزيع الطبيعى القياسى هو توزيع طبيعى وسطه الحسابى 0 وانحرافه المعيارى 1 (أى أن $\mu = 0$ و $\sigma = 1$) . ويمكن تحويل أى توزيع طبيعى (بوحدات X فى شكل ٣ - ٤) إلى توزيع طبيعى قياسى (بوحدات z) . وتحت هذه الشروط ، فإن 68.26% من المساحة (الاحتمال) تحت المنحنى الطبيعى القياسى تقع بين إحداثيين رأسيين يبعدان بمقدار انحراف معيارى واحد عن الوسط الحسابى (أى داخل $\mu \pm 1\sigma$) ، 95.54% تقع بين $\mu \pm 2\sigma$ ، 99.74 % تقع بين $\mu \pm 3\sigma$

ولإيجاد الاحتمالات (المساحات) فى مسائل تحتوى على التوزيع الطبيعى ، فإننا نحول أولا قيم X إلى قيم z المناظرة لها ، كالآتى :

$$z = \frac{X - \mu}{\sigma} \qquad (٣-١٤)$$



شكل ٣ - ٤

ثم نكشف عن قيمة z فى ملحق ٣ . ويعطى هذا قيمة الجزء من المساحة (الاحتمال) تحت المنحنى بين قيمة الوسط الحسابى وقيمة z

مثال (١٢) : المساحة (الاحتمال) تحت المنحنى الطبيعى المعيارى بين $z = 0$ و $z = 1.96$ نحصل عليها مقابلة للقيمة 1.96 فى ملحق ٣ . فى عمود z نبدأ بالقيمة 1.9 ونتحرك فى الصف المناظر لها حتى نصل إلى العمود المعنون 06. ، وتكون القيمة التى نحصل عليها 0.4750 . ويعنى هذا أن 47.50% من المساحة الكلية (1 أو 100 %) تحت المنحنى تقع بين $z = 0$ و $z = 1.96$ (المساحة المظللة فى الشكل فوق الجدول) . ولأن التوزيع متماثل ، فإن المساحة بين $z = 0$ و $z = -1.96$ (ليست مدرجة فى الجدول) هى أيضاً 0.4750 أو 47.50 % .

مثال (١٣) : افترض أن X متغير عشوائى يتبع التوزيع الطبيعى حيث $\mu = 10$ و $\sigma^2 = 4$ ونريد إيجاد احتمال أن تأخذ X قيمة بين 8 و 12 نحسب أولا قيمة z المناظرة لقيم X وهى 0 و 12 ثم نكشف عن القيم التى تناظر قيم z فى ملحق ٣ :

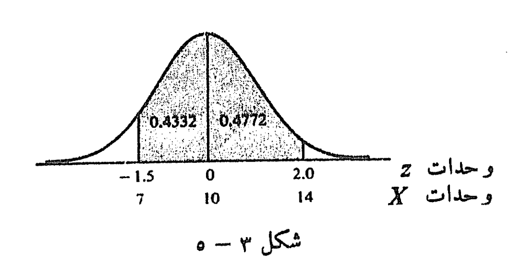
$$z_1 = \frac{X_1 - \mu}{\sigma} = \frac{8 - 10}{2} = -1 \qquad \text{و} \qquad z_2 = \frac{X_2 - \mu}{\sigma} = \frac{12 - 10}{2} = +1$$

عندما $z = 1$ نحصل على القيمة 0.3413 من ملحق (٣) . ومن ثم فإن المساحة بين $z = \pm 1$ تساوى $2(0.3413)$ أو 0.6826 . وهذا يعنى أن احتمال أن تأخذ X قيمة بين 8 و 12 أو $P(8 < X < 12)$ هو 68.26 % (انظر شكل ٣ – ٤)

مثال (١٤) : افترض مرة أخرى أن X متغير عشوائى موزع طبقاً للتوزيع الطبيعى حيث $\mu = 10$ ، و $\sigma^2 = 4$. فيمكن إيجاد احتمال أن X تأخذ قيمة بين 7 و 14 كالآتى :

$$z_1 = \frac{X_1 - \mu}{\sigma} = \frac{7 - 10}{2} = -1.5 \qquad \text{و} \qquad z_2 = \frac{X_2 - \mu}{\sigma} = \frac{14 - 10}{2} = 2$$

عندما $z_1 = 1.50$ ، نكشف عن القيمة 1.50 فى ملحق (٣) فنحصل على 0.4332 . وعندما $z_2 = 2$ نحصل على القيمة 0.4772 ومن ثم فإن $P = (7 < X < 14) = 0.4332 + 0.4772 = 0.9104$ أو 91.04 % (انظر شكل ٣ – ٥) . ومن ثم ، فإن احتمال أن X تأخذ قيمة أقل من 7 أو أكبر من 14 (المناطق غير المظللة فى أطراف التوزيع فى شكل ٣ – ٥) هى $1 - 0.91404 = 0.0896$ أى 8.96 % . ويستخدم التوزيع الطبيعى كتقريب للتوزيع ذى الحدين عندما $n \geq 30$ و $np > 5$ ، $n(1-p) > 5$ ، كما يستخدم كتقريب لتوزيع بواسون عندما $\lambda \geq 10$ (انظر المسائل ٣ – ٣٧ و ٣ – ٣٨) . وهناك توزيع احتمالى متصل آخر هو **التوزيع الأسى** (انظر المسألة ٣ – ٣٩) . وتنص **نظرية أو متباينة تشيبتشيف** أنه بصرف النظر عن شكل التوزيع ، فإن نسبة المشاهدات أو المساحة التى تقع فى حدود K انحراف معيارى على جانبى الوسط الحسابى تكون على الأقل $1 - 1/K^2$ ، حيث $K \geq 1$.
(انظر المسائل ٣ – ٤٠ و ٣ – ٧٢) .



شكل ٣ – ٥

مسائل محلولة

احتمال حدث منفرد :

٣ – ١ (أ) فرق بين الاحتمال الكلاسيكى أو المسبق ، وبين التكرار النسبى أو الاحتمال التجريبى وبين الاحتمال الشخصى أو الذاتى (ب) ماهى عيوب كل منها (جـ) لماذا ندرس نظرية الاحتمال ؟

(أ) طبقاً للاحتمال الكلاسيكى ، فإن احتمال حدث ما A هو :

$$P(A) = \frac{n_A}{N}$$

حيث $P(A) =$ احتمال وقوع الحدث A

$n_A =$ عدد الطرق التى يمكن أن يقع بها الحدث A

$N =$ عدد النواتج الكلية وكلها متساوية الفرصة فى الوقوع .

فى المدخل الكلاسيكى ، يمكن عمل تقارير احتمالية عن عملات متوازنة ، وأحجار نرد غير متحيزة وأوراق لعب بدون رمى عملة ، أو إلقاء نردة ، أو سحب ورقة من مجموعة أوراق لعب . أما التكرار النسبى أو الاحتمال التجريبى فيمثل النسبة بين عدد المرات التى يقع فيها حدث ما فعلا وبين العدد الكلى للنواتج الفعلية أو المشاهدات . وكلما زاد عدد المحاولات أو التجارب (مثل عدد مرات رمى العملة) . كلما اقترب التكرار النسبى أو الاحتمال التجريبى من الاحتمال الكلاسيكى أو المسبق. أما الاحتمال الشخصى أو الذاتى فيشير إلى درجة الاعتقاد لدى شخص ما بأن حدثاً معيناً سوف يحدث ، تأسيساً على أى أدلة متوافرة لديه .

(ب) الاحتمال الكلاسيكى أو المسبق يمكن تطبيقه فقط على ألعاب الصدفة (مثل رمى عملة متوازنة ، إلقاء حجر نرد غير متحيز ، أو سحب ورقة من مجموعة أوراق لعب عادية) حيث يمكننا أن نحدد مقدماً ، أى بدون تجريب ، احتمال حدوث حدث ما . وعادة فى مشاكل الاقتصاد والأعمال الحقيقية لايمكننا غالباً تعيين الاحتمالات مقدماً ومن ثم لايمكننا استخدام المدخل الكلاسيكى . ويتغلب التكرار النسبى أو الاحتمال التجريبى على عيوب المدخل الكلاسيكى باستخدام التكرارات النسبية التى حدثت فى الماضى كاحتمالات . ولكن الصعوبة فى التكرار النسبى أو المدخل التجريبى أننا نحصل على احتمالات مختلفة (تكرارات نسبية) عندما يختلف عدد المحاولات أو التجارب . وتستقر هذه الاحتمالات ، أو تقترب من نهاية ، مع زيادة عدد التجارب أو المحاولات . وحيث أن هذا قد يكون مكلفاً ويستغرق وقتاً ، فقد يلجأ الناس إلى استخدامه بدون عدد « كاف » من التجارب أو المحاولات . أما عيب المدخل الشخصى أو الذاتى فهو أن الأفراد المختلفين قد يعطون احتمالات مختلفة تماماً عندما يجابهون بنفس الموقف .

(جـ) معظم القرارات التى نجابهها فى الاقتصاد ، والإدارة ، والعلوم ، وفى حياتنا اليومية تتضمن مخاطرة واحتمالات . هذه الاحتمالات يمكن فهمها وتوضيحها بشكل أسهل فى مباريات الاختيار لأنه يمكن تعيين احتمالات موضوعية للأحداث المختلفة بسهولة فى مثل هذه الحالات . ومع ذلك فإن السبب الرئيسى لدراسة نظرية الاحتمالات هو المساعدة على اتخاذ قرارات ذكية فى الاقتصاد ، والأعمال ، والعلوم ، والحياة اليومية عندما تتضمن مخاطرة أو عدم تأكد .

٣ — ٢ ماهو احتمال (أ) صورة فى رمية واحدة لعملة متوازنة ؟ كتابة ؟ صورة أو كتابة ؟ (ب) ظهور العدد 2 فى رمية واحدة لنردة غير متحيزة ؟ عدد غير 2 ؟ عدد 2 أو عدد ليس 2 ؟

(أ)

$$P(\mathrm{H}) = \frac{n_\mathrm{H}}{N} = \frac{1}{2}$$

$$P(\mathrm{T}) = \frac{n_\mathrm{T}}{N} = \frac{1}{2}$$

$$P(\mathrm{H}) + P(\mathrm{T}) = \frac{1}{2} + \frac{1}{2} = 1$$

(ب) حيث أن كلا من الوجوه الستة لنردة غير متحيزة لها نفس الإمكانية فى الظهور وأن 2 هى إحدى هذه الامكانيات فإن ،

$$P(2) = \frac{n_2}{N} = \frac{1}{6}$$

احتمال عدم الحصول على 2 (أى $P(2')$) هو

$$P(2') = 1 - P(2) = 1 - \frac{1}{6} = \frac{5}{6}$$

$$P(2) + P(2') = \frac{1}{6} + \frac{5}{6} = \frac{6}{6} = 1$$ (أو تأكد تام)

٣ — ٣ عند سحب ورقة واحدة من مجموعة أوراق لعب مخلوطة جيداً ما احتمال أن تكون الورقة (أ) شايب (ب) بستونى (جـ) الشايب البستونى (د) ليست الشايب البستونى أو (هـ) الشايب البستونى أو ليس الشايب البستونى ؟

(أ) حيث أن هناك 4 شايب K ، فى المجموعة العادية المكونة من 52 ورقة فإن

$$P(K) = \frac{n_K}{N} = \frac{4}{52} = \frac{1}{13}$$

(ب) حيث أن هناك 13 ورقة بستونى S ، من إجمالى 52 ورقة ، $P(S) = 13/52 = 1/4$

(جـ) هناك شايب بستونى واحد فى مجموعة أوراق اللعب فيكون $P(K_S) = 1/52$

(د) احتمال عدم سحب الشايب البستونى $(P(K'_S) = 1 - 1/52 = 51/52$

(هـ) $P(K_S) + P(K'_S) = 1/52 + 51/52 = 52/52 = 1$ ، أى تأكد تام

٣ — ٤ وعاء يحتوى على 10 كرات متماثلة تماماً باستثناء أن 5 منها حمراء ، 3 زرقاء و 2 خضراء . سحبت كرة من الوعاء . ماهو احتمال أن تكون (أ) حمراء ؟ (ب) زرقاء (جـ) خضراء ؟ (د) ليست زرقاء (هـ) ليست خضراء (و) خضراء أوليست خضراء ؟ (ز) ماهو معامل الترجيح لصالح سحب كرة زرقاء ؟ (ح) ماهو معامل الترجيح لصالح سحب كرة غير زرقاء ؟

(أ)
$$P(R) = \frac{n_R}{N} = \frac{5}{10} = 0.5$$

(ب)
$$P(B) = \frac{n_B}{N} = \frac{3}{10} = 0.3$$

(جـ)
$$P(G) = \frac{n_G}{N} = \frac{2}{10} = 0.2$$

(د)
$$P(B') = 1 - P(B) = 1 - 0.3 = 0.7$$

(هـ)
$$P(G') = 1 - P(G) = 1 - 0.2 = 0.8$$

(و)
$$P(G) + P(G') = 0.2 + 0.8 = 1$$

(ز) معامل الترجيح لصالح التقاط كرة زرقاء يعبر عن النسبة بين عدد طرق التقاط كرة زرقاء وعدد طرق التقاط كرة غير زرقاء ، وحيث أن هناك 3 كرة زرقاء و 7 كرة غير زرقاء ، فإن معامل الترجيح لصالح التقاط كرة زرقاء هو 3 إلى 7 أو 3:7 .

(ح) معامل الترجيح لصالح التقاط كرة غير زرقاء هو 7 إلى 3 أو 7:3

٣ — ٥ افترض أن الرقم 3 ظهر 106 مرة فى 600 رمية لنردة (أ) ما هو التكرار النسبى للرقم 3 ؟ كيف يختلف هذا عن الاحتمال الكلاسيكى أو المسبق ؟ (ب) إذا زاد عدد مرات رمى النردة ماذا تتوقع أن يكون التكرار النسبى أو الاحتمال التجريبى ؟

(أ) التكرار النسبى أو الاحتمال التجريبى للرقم 3 هو النسبة بين عدد مرات ظهور الرقم 3 (106) والعدد الكلى لمرات رمى النردة (600) . وعليه فإن التكرار النسبى أو الاحتمال التجريبى للرقم 3 هو $106/600 \cong 0.177$ فى 600 رمية للنردة . طبقاً للاحتمال الكلاسيكى أو المسبق (بدون رمى نردة على الإطلاق) ، $P(3) = 1/6 \cong 0.167$ فإذا كانت النردة غير متحيزة فإننا نتوقع أن الرقم 3 يظهر 100 مرة من بين 600 رمية للنردة بالمقارنة بالعدد الفعلى ، المشاهد ، أو التجريبى 106 مرة .

(ب) إذا زاد عدد مرات رمى نفس الزهرة عن 600 ، فإننا نتوقع أن يزداد التكرار النسبى أو الاحتمال التجريبى اقتراباً من الاحتمال الكلاسيكى أو المسبق .

٣ - ٦ عملية إنتاجية ينتج عنها 27 وحدة معيبة فى كل 1000 وحدة منتجة (أ) ماهو التكرار النسبى أو الاحتمال التجريبى للوحدة المعيبة ؟ (ب) من بين إنتاج يومى قدره 1600 وحدة ، كم عدد الوحدات المعيبة التى تتوقعها ؟

(أ) التكرار النسبى أو الاحتمال التجريبى لوحدة المعيبة هو $27/1,000 = 0.027$

(ب) بضرب عدد الوحدات المنتجة يومياً فى التكرار النسبى أو الاحتمال التجريبى للوحدة المعيبة (0.027) ، نحصل على عدد الوحدات المعيبة المتوقعة بين الإنتاج اليومى . ويكون هذا $(1600)(0.027) = 43$ لأقرب وحدة .

احتمال الأحداث المتعددة :

٣ - ٧ عرف واعط بعض الأمثلة على الأحداث التى تكون (أ) متنافية الحدوث (ب) ليست متنافية الحدوث (ج) مستقلة ، (د) غير مستقلة .

(أ) يعتبر حدثان أو أكثر متنافيين أو غير مشتركين ، إذا كان وقوع أحدهما يحجب وقوع الآخر (الأخرى) . فإذا وقع واحد من الأحداث فإن الآخر (الأخرى) لا يقع . فمثلا ، عند رمى عملة حرة واحدة نحصل إما على صورة وإما على كتابة ولكن ليس الإثنان معاً . ومن ثم فإن الصور والكتابة حدثان متنافيان . وفى رمية واحدة لنردة فإننا نحصل على واحد وواحد فقط من ستة نواتج ممكنة : 1, 2, 3,4, 5, أو 6 . فالنواتج هذه متنافية الحدوث . وورقة اللعب المسحوبة عشوائياً يمكن أن تكون من نوع واحد فقط : دينارى ، كوبة ، سباتى ، أو بستونى . الطفل المولود يكون إما ولدا وإما بنتاً . والوحدة المنتجة على خط الإنتاج هى إما جيدة وإما معيبة .

(ب) يعتبر حدثان (أو أكثر) غير متنافين إذا كان من الممكن حدوثهما معاً . فحدوث أحدهما لايحجب حدوث الآخر (الأخرى) . فعلى سبيل المثال ورقة اللعب المسحوبة عشوائياً من الممكن أن تكون آس وسباتى فى نفس الوقت ، ومن ثم فإن الآس والسباتى ليسا حدثين متنافيين بالتبادل إذا يمكننا سحب الآس السباتى . وحيث أنه من الممكن أن يحدث تضخم وكساد فى آن واحد ، فإن التضخم والكساد ليسا حدثين متنافيين .

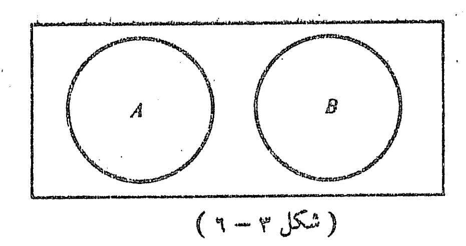
(ج) يكون حدثان (أو أكثر) مستقلين إذا كان وقوع أحدهما لايؤثر على أى نحو فى وقوع الآخر (الأخرى) . فعلى سبيل المثال ، عند رمى عملة مرتين متتاليتين ، فإن ناتج الرمية الثانية لايعتمد على ناتج الرمية الأولى . وهذا ينطبق أيضاً على رميتين متتاليتين لنردة أو سحب ورقتين من مجموعة أوراق لعب إذا كان هناك إعادة الورقة المسحوبة .

(د) يعتبر حدثان (أو أكثر) غير مستقلين إذا كان يؤثر أحدهما على احتمال وقوع الآخر (الأخرى) . فعلى سبيل المثال ، إذا سحبنا ورقة من مجموعة أوراق لعب ولم نعدها إليها ، فإن احتمال سحب نفس الورقة فى السحب الثانى هو 0 . كما أن الاحتمالات الأخرى تتأثر كلها ، حيث أصبح فى المجموعة الآن 51 ورقة فقط . وبالمثل ، إذا كانت نسبة المعيب فى وردية المساء أعلى منها فى وردية الصباح ، فإن احتمال أن وحدة مسحوبة من إنتاج المساء تكون معيبة أعلى منه بالنسبة لإنتاج الصباح .

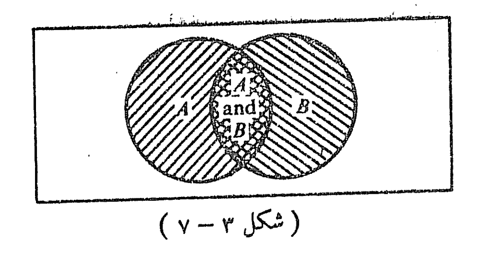
٣ - ٨ ارسم شكل فن (أ) للأحداث المتنافية (ب) للأحداث غير المتنافية (ج) هل الأحداث المتنافية مستقلة أم غير مستقلة ؟ لماذا ؟

(أ) شكل (٣ — ٦) يوضح شكل فن للحدثين A و B المتنافيين.

(ب) شكل (٣ — ٧) يوضح شكل فن للحدثين A و B غير المتنافيين



(شكل ٣ — ٦)



(شكل ٣ — ٧)

(جـ) الأحداث المتنافية أحداث غير مستقلة عندما يقع واحد من الأحداث . فاحتمال وقوع الآخر يكون 0 فوقوع الأول يؤثر على (يحجب) وقوع الثانى .

٣ — ٩ ماهو احتمال الحصول على : (أ) أقل من 3 فى رمية واحدة لنردة غير متحيزة ؟

(ب) كوبة أو سباتى عند سحب ورقة واحدة من مجموعة عادية مخلوطة خلطاً جيداً ؟

(جـ) كرة حمراء أو زرقاء من وعاء يحتوى على 5 كرات حمراء ، 3 زرقاء ، 2 خضراء ؟

(د) أكثر من 3 فى رمية واحدة لنردة غير متحيزة ؟

(أ) الحصول على أقل من 3 فى رمية واحدة (لنردة غير متحيزة يعنى الحصول على 1 أو 2).وهذه أحداث متنافية . بتطبيق قاعدة الجمع للأحداث المتنافية نحصل على

$$P(1 \text{ أو } 2) = P(1) + P(2) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$$

باستخدام نظرية المجموعات ، يمكن كتابة (1 أو 2) P فى صورة مكافئة كالآتى $P(1 \cup 2)$ وتقرأ $\cup$ « اتحاد » كبديل عن « أو » .

(ب) الحصول على كوبة أو سباتى عند سحب ورقة واحدة من مجموعة عادية مخلوطة تكون أحداثاً متنافية الحدوث أيضاً . بتطبيق قاعدة الجمع نحصل على

$$P(\text{H أو C}) = P(\text{H} \cup \text{C}) = \frac{13}{52} + \frac{13}{52} = \frac{26}{52} = \frac{1}{2}$$

(جـ)
$$P(\text{R أو B}) = P(\text{R} \cup \text{B}) = \frac{5}{10} + \frac{3}{10} = \frac{8}{10} = \frac{4}{5} = 0.8$$

(د)
$$P(4 \text{ أو } 5 \text{ أو } 6) = P(4 \cup 5 \cup 6) = P(4) + P(5) + P(6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}$$

٣ — ١٠ (أ) ماهو احتمال الحصول على آس أو سباتى عند سحب ورقة واحدة من مجموعة مخلوطة جيداً ؟ (فى كل المسائل الباقية ، سوف يكون من المفترض ضمنياً أن العملات متوازنة ، والنردات غير متحيزة وأن أوراق اللعب تسحب من مجموعة عادية مخلوطة جيداً وبدون إحلال) . (ب) ماهى وظيفة الحد السالب فى قاعدة الجمع ، للأحداث غير المتنافية ؟

(أ) الحصول على آس أو على سباتى لايكونان حدثين متنافيين لأنه يمكننا الحصول على الآس السباتى . بتطبيق قاعدة الجمع للأحداث غير المتنافية نحصل على

$$P(\text{A أو C}) = P(\text{A}) + P(\text{C}) - P(\text{A و C}) = \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52} = \frac{4}{13}$$

ويمكن كتابة عبارة الاحتمالات السابقة فى صورة مكافئة باستخدام نظرية المجموعات كالآتى :

$$P(A \cup C) = P(A) + P(C) - P(A \cap C)'$$

حيث $\cap$ تقرأ « تقاطع » وتستخدم بدلا من « و » .

(ب) وظيفة الحد السالب فى قاعدة الجمع للأحداث غير المتنافية هى تجنب العد المزدوج . فعلى سبيل المثال ، عند حساب $P(A$ أو $C)$ فى الجزء (أ) فإن الآس السباتى يتم عده مرتين ، مرة كآس ومرة كسباتى . ومن ثم فإننا نطرح احتمال الحصول على الآس السباتى لتجنب هذا الحساب المزدوج أما إذا كان الحدثان متنافيين فإن احتمال حدوثهما معاً فى آن واحد يكون 0 ، ولا يوجد أى حساب مزدوج . ولهذا فإن قاعدة الجمع للأحداث المتنافية لاتتضمن حدا سالباً .

٣ - ١١ ماهو احتمال (أ) تضخم I أو كساد R ، إذا كان احتمال التضخم 0.3 واحتمال الكساد 0.2 واحتمال التضخم والكساد 0.06 ؟ (ب) سحب آس ، سباتى ، أو دينارى عند سحب ورقة واحدة من مجموعة أوراق لعب ؟

(أ) حيث أن احتمال تضخم وكساد معاً ليس 0 ، فالتضخم والكساد ليسا حدثين متنافيين بتطبيق قاعدة الجمع ، نحصل على

$$P(I \text{ أو } R) = P(I) + P(R) - P(I \text{ و } R)$$
$$P(I \cup R) = P(I) + P(R) - P(I \cap R)$$
$$P(I \text{ أو } R) = P(I \cup R) = 0.3 + 0.2 - 0.06 = 0.44$$

(ب) الحصول على آس ، سباتى ، أو دينارى لايشكل أحداثاً متنافية لأنه يمكننا الحصول على الآس السباتى أو الآس الدينارى بتطبيق قاعدة الجمع للاحداث غير المتنافية نحصل على

$$P(A \text{ أو } C \text{ أو } D) = P(A) + P(C) + P(D) - P(A \text{ و } C) - P(A \text{ و } D)$$
$$P(A \text{ أو } C \text{ أو } D) = \frac{4}{52} + \frac{13}{52} + \frac{13}{52} - \frac{1}{52} - \frac{1}{52} = \frac{28}{52} = \frac{7}{13}$$

٣ - ١٢ ماهو احتمال الحصول : (أ) 6 و 6 فى رميتين لحجر نرد ؟ (ب) 6 على كل نردة عند رمى نردتين ؟ (جـ) كرتين زرقاوين فى سحبين متتاليين مع الإحلال من الوعاء فى مسألة (٣ - ٤) ؟ (د) ثلاث بنات فى عائلة لديها 3 أطفال ؟

(أ) الحصول على 6 فى كل من الرميتين للنردة يمثل حدثين مستقلين . بتطبيق قاعدة الضرب للأحداث المستقلة ، نحصل على

$$P(6 \text{ و } 6) = P(6 \cap 6) = P(6) \cdot P(6) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

(ب) الحصول على 6 فى كل نردة فى رمية واحدة لها يمثل حدثين أيضاً حدثين مستقلين . فيكون

$$P(6 \text{ و } 6) = P(6 \cap 6) = P(6) \cdot P(6) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

(جـ) حيث أننا نعيد الكرة الأولى بعد سحبها ، فإن احتمال الحصول على كرة زرقاء فى السحب الثانى يكون هو نفسه كما فى السحب الأول . الأحداث مستقلة ، وعليه ،

$$P(B \text{ و } B) = P(B \cap B) = P(B) \cdot P(B) = \frac{3}{10} \cdot \frac{3}{10} = \frac{9}{100} = 0.09$$

(د) إن حدث بنت ، G ، عند كل ميلاد ، يشكل حدث مستقل ، وحيث أن احتمال البنت فى كل ميلاد هو 0.5 ، فإن

$$P(G \text{ و } G \text{ و } G) = P(G \cap G \cap G) = P(G) \cdot P(G) \cdot P(G) = (0.5) \cdot (0.5) \cdot (0.5) = 0.125$$

ى فرصة واحدة لكل 8 حالات

٣ - ١٣ (أ) اسرد كل النواتج الممكنة لإلقاء نردتين فى آن واحد (ب) ماهو احتمال الحصول على مجموع 5 عند إلقاء نردتين فى آن واحد ؟ (جـ) ماهو احتمال الحصول على مجموع 4 أو أقل عند رمى نردتين فى آن واحد ؟ أكثر من 4 ؟

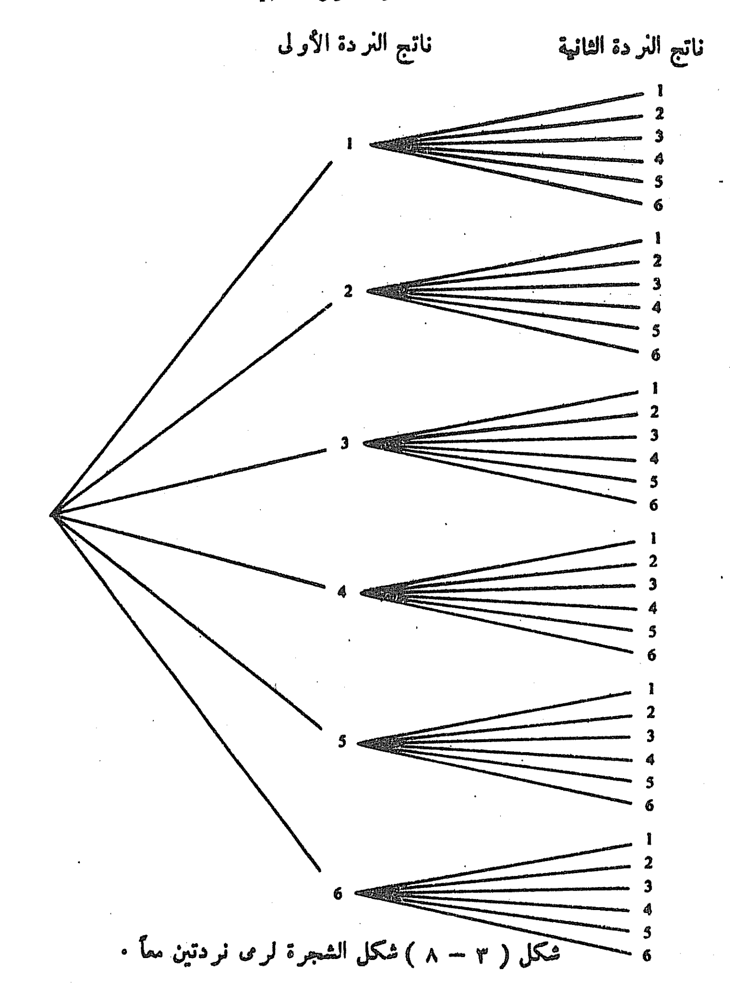
(أ) كل نردة لها 6 نواتج متساوية الامكان ، وناتج كل نردة مستقل . وحيث أن كلا من النواتج الستة للنردة الأولى ، يمكن أن يظهر مع أى من النواتج الستة للنردة الثانية ، فإن هناك 36 ناتجاً ممكناً للاثنين معاً . أى أن ، بفضاء العينة ، N ، 36 عنصراً . (فى جدول (٣-٢)) ، العدد الأول يشير إلى ناتج النردة الأولى ، والعدد الثانى يشير إلى ناتج النردة الثانية ويمكن تمييز النردتين باستخدام لونين مختلفين) ويمكن توضيح العدد الكلى المكون من 36 ناتجاً ممكناً باستخدام شكل الشجرة أو الشكل التتابعى كما فى شكل (٣ - ٨) .

جدول (٣ - ٢) نواتج رمى نردتين معاً

1, 1	2, 1	3, 1	4, 1	5, 1	6, 1
1, 2	2, 2	3, 2	4, 2	5, 2	6, 2
1, 3	2, 3	3, 3	4, 3	5, 3	6, 3
1, 4	2, 4	3, 4	4, 4	5, 4	6, 4
1, 5	2, 5	3, 5	4, 5	5, 5	6, 5
1, 6	2, 6	3, 6	4, 6	5, 6	6, 6

(ب) من بين 36 ناتجاً ممكناً ومتساوى الفرصة فى الحدوث ، هناك 4 نواتج تعطى مجموعاً قدره 5 . وهذه هى ، 4, 1 ، و3,2 ، 2,3 ، 1,4 وعليه فإن احتمال مجموع 5 (حدث A) عند رمى نردتين معاً يكون

$$P(A) = \frac{n_A}{N} = \frac{4}{36} = \frac{1}{9}$$



شكل (٣ - ٨) شكل الشجرة لرمى نردتين معاً .

(جـ) الحصول على مجموع 4 أو أقل يتضمن الحصول على مجموع 2 ، 3 ، او 4 . هناك 6 نواتج مجموعها 4 أو أقل وهذه هى 1, 1 ، 1, 2 ، 1, 3 ، 2, 1 ، 2, 2 ، 3, 1 . ومن ثم إذا عرفنا الحدث A بأنه الحصول على مجموع 4 أو أقل ، $P(A) = 6/36 = 1/6$ واحتمال الحصول على مجموع أكثر من 4 يساوى 1 ناقصاً احتمال الحصول على مجموع 4 أو أقل . ويكون هذا $1 - 1/6 = 5/6$.

٣ – ١٤ ماهو احتمال (أ) التقاط كرة حمراء ثانية من الوعاء فى المسألة (٣ – ٤) علماً بأنه قد تم التقاط كرة حمراء فى المرة الأولى ولم تعد إلى الوعاء ؟ (ب) كرة حمراء فى المرة الثانية علماً بأن الأولى لم تكن حمراء ولم تعد إلى الوعاء ؟ (جـ) كرة حمراء فى المرة الثالثة عندما نكون قد حصلنا على كرة حمراء وكرة غير حمراء فى المرتين الأولى والثانية ولم تعادا ؟

(أ) سحب كرة حمراء ثانية عندما تكون الأولى حمراء ولم تعد هو حدث غير مستقل ، حيث أصبح هناك عدد 4 كرة حمراء و 5 كرة غير حمراء باقية فى الوعاء . **الاحتمال الشرطى** أن الكرة الثانية حمراء بعد ماتم سحب كرة حمراء فى المرة الأولى ولم تعد هو $P(R/R) = 4/9$

(ب) الاحتمال الشرطى للحصول على كرة حمراء فى المرة الثانية بعد ما تم سحب كرة غير حمراء R' فى المرة الأولى ولم تعد إلى الوعاء هو $P(R/R') = 5/9$

(جـ) حيث أن كرتين إحداهما ليست حمراء تم سحبهما ولم تعادا ، فإنه يبقى 8 كرات فى الوعاء منها 4 حمراء . الاحتمال الشرطى لالتقاط كرة حمراء اخرى هو $P(R/R \text{ و } R') = P(R/R' \text{ و } R) = 4/8 = 1/2$

٣ – ١٥ ماهو احتمال الحصول على (أ) كرتين حمراوين من الوعاء فى المسألة (٣ – ٤) عند السحب مرتين بدون إعادة ؟ (ب) 2 آس عند سحب ورقتى مجموعة لعب بدون إحلال؟ (جـ) آس سباتى وبستونى على الترتيب عند سحب ورقتى مجموعة لعب بدون إحلال ؟ (د) بستونى وآس سباتى على الترتيب عند سحب ورقتى مجموعة لعب بدون إحلال ؟ (هـ) ثلاث كرات حمراء من الوعاء فى مسألة (٣ – ٤) عند السحب 3 مرات بدون إحلال ؟ (و) ثلاث كرات حمراء عند السحب ثلاث مرات مع الإحلال ؟

(أ) بتطبيق قاعدة الضرب للأحداث غير المستقلة ، نحصل على

$$P(R \text{ و } R) = P(R \cap R) = P(R) \cdot P(R/R) = \frac{5}{10} \cdot \frac{4}{9} = \frac{20}{90} = \frac{2}{9}$$

(ب)
$$P(A \text{ و } A) = P(A \cap A) = P(A) \cdot P(A/A) = \frac{4}{52} \cdot \frac{3}{51} = \frac{12}{2{,}652} = \frac{1}{221}$$

(جـ)
$$P(A_C \text{ و } S) = P(A_C \cap S) = P(A_C) \cdot P(S/A_C) = \frac{1}{52} \cdot \frac{13}{51} = \frac{13}{2{,}652}$$

(د)
$$P(S \text{ و } A_C) = P(S \cap A_C) = P(S) \cdot P(A_C/S) = \frac{13}{52} \cdot \frac{1}{51} = \frac{13}{2{,}652} = P(A_C \text{ و } S)$$

(هـ)
$$P(R \text{ و } R \text{ و } R) = P(R \cap R \cap R) = P(R) \cdot P(R/R) \cdot P(R/R \text{ و } R)$$
$$= \frac{5}{10} \cdot \frac{4}{9} \cdot \frac{3}{8} = \frac{60}{720} = \frac{1}{12}$$

(و) مع الإحلال ، فإن سحب ثلاث كرات حمراء من وعاء يشكل أحداثاً مستقلة . وعليه فإن ،

$$P(R \text{ و } R \text{ و } R) = P(R) \cdot P(R) \cdot P(R) = \frac{5}{10} \cdot \frac{5}{10} \cdot \frac{5}{10} = \frac{125}{1{,}000} = \frac{1}{8} = 0.125$$

٣ – ١٦ أظهرت التجارب الماضية انه من بين كل 100,000 وحدة منتجة فى الوردية الصباحية هناك 200 وحدة معيبة ، وأنه من بين كل 100,000 وحدة منتجة فى الوردية المسائية ، هناك 600 وحدة معيبة . خلال كل 24 ساعة ، هناك 1,000

وحدة منتجة فى الوردية الصباحية و 600 وحدة منتجة فى الوردية المسائية . ماهو احتمال أن تكون وحدة مسحوبة من الإنتاج اليومى الكلى وقدره 600 وحدة (أ) من إنتاج الوردية الصباحية وأنها معيبة ؟ (ب) من إنتاج الوردية المسائية وأنها معيبة ؟ (جـ) من إنتاج الوردية المسائية وأنها ليست معيبة ؟ (د) أنها معيبة ، سواء أنتجت فى الوردية الصباحية أو المسائية ؟

(أ) احتمال أن تكون الوحدة المسحوبة من إنتاج الوردية الصباحية M ، ومن إنتاج الوردية المسائية E هو

$$P(\mathrm{M}) = \frac{1{,}000}{1{,}600} = 0.625 \qquad و \qquad P(\mathrm{E}) = \frac{600}{1{,}600} = 0.375$$

احتمال سحب وحدة معيبة D من إنتاج الوردية الصباحية ومن إنتاج الوردية المسائية هما على الترتيب

$$P(\mathrm{D/M}) = \frac{200}{100{,}000} = 0.002 \qquad و \qquad P(\mathrm{D/E}) = \frac{500}{100{,}000} = 0.005$$

احتمال أن تكون الوحدة المسحوبة عشوائياً من إجمالى الإنتاج اليومى هى من إنتاج الوردية الصباحية وأنها معيبة هو

$$P(\mathrm{M} \text{ و } \mathrm{D}) = P(\mathrm{M}) \cdot P(\mathrm{D/M}) = (0.625)(0.002) = 0.00125$$

(ب)

$$P(\mathrm{E} \text{ و } \mathrm{D}) = P(\mathrm{E}) \cdot P(\mathrm{D/E}) = (0.375)(0.005) = 0.001875$$

(جـ)

$$P(\mathrm{E} \text{ و } \mathrm{D'}) = P(\mathrm{E}) \cdot P(\mathrm{D'/E}) = (0.375)\frac{99{,}500}{100{,}000} = 0.373125$$

(د) عدد الوحدات المعيبة المتوقع فى وردية الصباح يساوى احتمال الوحدة المعيبة فى إنتاج الصباح مضروباً فى عدد الوحدات المنتجة فى وردية الصباح . ويكون هذا 2 = (1, 000) (0. 002) ومن وردية المساء نتوقع عدد 3 = (600) (0.005) وحدة معيبة . أى أننا نتوقع 5 وحدة معيبة من إجمالى 1600 وحدة منتجة خلال فترة 24 ساعة . فإذا كان هناك 5 وحدة معيبة ، فإن احتمال سحب أى منها عشوائياً من إجمالى عدد الوحدات وقدرها 1600 وحدة يكون 5/1, 600 أو 1/320 أو 0.003125 .

٣ – ١٧ (أ) من قاعدة الضرب للأحداث غير المستقلة B و A ، اشتق صيغة $P(A/B)$ بدلالة $P(B/A)$ و $P(B)$. تعرف هذه الصيغة باسم نظرية بييز وتستخدم لتعديل الاحتمالات عندما تتاح بيانات إضافية (ب) باستخدام نظرية بييز ، أوجد احتمال أن وحدة معيبة تم سحبها من الإنتاج اليومى وقدره 1600 وحدة فى مسألة (٣ – ١٦) ، كانت من إنتاج الوردية الصباحية ، واحتمال أنها كانت من إنتاج اوردية المسائية .

(أ)

$$P(B \text{ و } A) = P(B) \cdot P(A/B)$$

بقسمة الطرفين على $P(B)$ وإعادة الترتيب ، نحصل على

$$P(A/B) = \frac{P(B \text{ و } A)}{P(B)}$$

ولكن ، $P(B \text{ و } A) = P(A \text{ و } B)$ انظر مسألة (٣ – ١٥) (جـ) و (د) . فيكون

(٣ – ١٥) نظرية بييز

$$P(A/B) = \frac{P(A \text{ and } B)}{P(B)} \qquad و \qquad P(A/B) = \frac{P(A) \cdot P(B/A)}{P(B)}$$

(ب) بتطبيق نظرية بييز على مسألة (٣ – ١٦) ، وباستخدام M للدلالة على الوردية الصباحية بدلا من A و D للدلالة على الوحدات معيبة بدلا من B وباستخدام نتائج المسألة (٣ – ١٦) ، نحصل على

$$P(\mathrm{M/D}) = \frac{P(\mathrm{M}) \cdot P(\mathrm{D/M})}{P(\mathrm{D})} = \frac{(0.625)(0.002)}{0.003125} = \frac{0.00125}{0.003125} = 0.4$$

أى أنه إذا سحبت وحدة معيبة عشوائياً من إجمالى الإنتاج اليومى وقدره 1600 وحدة ، فاحتمال أنها من وردية الصباح هو 40% وبالمثل ،

$$P(E/D) = P(E) \cdot P(D/E) = \frac{(0.375)(0.005)}{0.003125} = \frac{0.001875}{0.003125} = 0.6, \text{ أو } 60\%$$

ويمكن تعميم نظرية بيز ، فمثلا يمكن إيجاد احتمال وحدة معيبة ، B ، قد التقطت عشوائيا هى من إنتاج أى من n مصنع ($A_i, i = 1, 2, \dots n$) كالآتى :

$$P(A_i/B) = \frac{P(A_i) \cdot P(B/A_i)}{\sum P(A_i) \cdot P(B/A_i)} \qquad (٣-١٦)$$

حيث تشير Σ إلى المجموع لعدد n مصنع (التى تنتج المنتج) . تطبق نظرية بيز فى اتخاذ القرارات فى قطاع الأعمال ، ولكنها ليست مستخدمة كثيراً فى مجال الاقتصاد (وإن كان الاقتصاد القياسى المؤسس على نظرية بيز تتزايد أهميته) .

٣ — ١٨ ناد به 8 أعضاء . (أ) كم عدد اللجان المختلفة المكونة من 3 أعضاء يمكن تكوينها فى النادى ؟ (تختلف لجنة عن أخرى حتى لو اختلف عضو واحد فقط) (ب) كم عدد اللجان المختلفة المكونة من 3 أعضاء يمكن تشكيلها فى النادى إذا كان لكل لجنة مختارة رئيس وأمين صندوق وسكرتير ؟ .

(أ) يهمنا هنا إيجاد عدد التوافيق المختلفة التى يمكن عملها من 8 أفراد إذا أخذوا ثلاثة فى كل مرة بصرف النظر عن الترتيب .

$${}_8C_3 = \frac{8!}{3!\,(8-3)!} = \frac{8!}{3!\,5!} = \frac{8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{3 \cdot 2 \cdot 1 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = \frac{8 \cdot 7 \cdot 6}{3 \cdot 2 \cdot 1} = \frac{336}{6} = 56$$

وعموماً فإن عدد تكوينات n من الأشياء مأخوذة X من الأشياء فى كل مرة بصرف النظر عن الترتيب هو توفيق ويحسب كالآتى :

$${}_nC_X = \binom{n}{X} = \frac{n!}{X!\,(n-X)!} \qquad (٣-١٧)$$

حيث $n!$ (وتقرأ مضروب n) $= n.(n-1).(n-2) \dots 3.2.1$ و $0! = 1$ بالتعريف

(ب) حيث أن كل لجنة من 3 أفراد تتقاسم وظائف رئيس ، وأمين صندوق وسكرتير ، فإن مايهمنا الآن إيجاد عدد التباديل لعدد 8 أفراد مأخوذاً 3 أفراد فى كل مرة ، والترتيب هنا مهم :

$${}_8P_3 = \frac{8!}{(8-3)!} = \frac{8!}{5!} = \frac{8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1} = 8 \cdot 7 \cdot 6 = 336$$

وبصفة عامة ، فإن عدد التكوينات ، فى ترتيب محدد ، لعدد n من الأشياء مأخوذة X من الأشياء فى كل مرة هو تبديل ويحسب كالآتى :

$${}_nP_X = \frac{n!}{(n-X)!} \qquad (٣-١٨)$$

والتباديل والتوافيق (ويشار إليها عادة كأساليب العد) تساعد فى حساب عدد المرات المتساوية الإمكان لحدوث الحدث A بالنسبة إلى عدد النواتج المتساوية الإمكان الكلية . ولم تستخدم التباديل والتوافيق فى المسائل السابقة لأن تلك المسائل كانت بسيطة .

التوزيعات الاحتمالية المنفصلة : التوزيع ذو الحدين :

٣ — ١٩ عرف المقصود بكل ما يأتى واعط مثالا (أ) متغير عشوائى (ب) متغير عشوائى منفصل ، (جـ) توزيع احتمالى منفصل ، (د) ماهو الفرق بين توزيع احتمالى وتوزيع التكرار النسبى ؟

(أ) المتغير العشوائى هو المتغير الذى ترتبط كل قيمة من قيمه باحتمال تحقق تلك القيمة . فعلى سبيل المثال عند رمى نردة ، هناك 6 نواتج متنافية (1, 2, 3, 4, 5, أو 6) ، لكل منها احتمال وقوع قدره 1/6 ومن ثم فإن ناتج رمى نردة هو متغير عشوائى .

(ب) المتغير العشوائى المنفصل هو المتغير الذى يمكن أن يأخذ فقط عدداً محدوداً أو متميزاً من القيم . فعلى سبيل المثال ، فإن ناتج رمى نردة يمثل متغيراً عشوائياً منفصلا لأن قيمه محدودة بالأعداد 1, 2, 3, 4, 5 أو 6 وهذا على عكس المتغيرات المتصلة والتى يمكن أن تأخذ عدداً غير نهائى من القيم داخل أى فترة معلومة . (انظر المسألة ٣ – ٣١ (أ)) .

(ﺟ) التوزيع الاحتمالى المنفصل يشير إلى مجموعة كل قيم متغير عشوائى (منفصل) والاحتمالات المناظرة لها . فالمجموعة المكونة من 6 نواتج لرمى نردة واحتمالاتها المناظرة هى مثال لتوزيع احتمالى منفصل . ويكون مجموع الاحتمالات المرتبطة بكل القيم التى يمكن أن يأخذها متغير عشوائى منفصل مساوياً 1 .

(د) التوزيع الاحتمالى يشير إلى الاحتمالات الكلاسيكية أو المسبقة المناظرة لكل القيم التى يمكن أن يأخذها متغير عشوائى . ولأنه يتم تعيين هذه الاحتمالات مقدماً وبدون أى تجارب ، فإن التوزيع الاحتمالى عادة مايشار إليه كتوزيع تكرارى نظرى (نسبى) . ويختلف هذا عن التوزيع التكرارى التجريبى (النسبى) ، والذى يشير إلى النسبة بين عدد المرات التى يحدث فيها ناتج معين فعلا إلى إجمالى عدد المحاولات الفعلية أو المشاهدات . فعلى سبيل المثال ، عندما نرى نردة عدداً من المرات فعلا ، فليس من المتوقع أن نحصل على كل ناتج سدس عدد مرات الرمى بالضبط. ولكن ، مع تزايد عدد المحاولات فإن التوزيع التكرارى التجريبى (النسبى) يستقر عند الاحتمال (المنتظم) أو توزيع التكرار – النسبى النظرى وقدره 1/6 .

٣ – ٢٠ اشتق صيغة (أ) الوسط الحسابى μ أو القيمة المتوقعة $E(X)$ ، (ب) التباين لتوزيع احتمالى منفصل .

(أ) معادلة الوسط الحسابى لبيانات المجتمع المبوبة (معادلة (٢ – ٢ أ)) هى

$$\mu = \frac{\sum fX}{N}$$

حيث ΣfX هى مجموع تكرار كل فئة f مضروباً فى مركز الفئة X وحيث $N = \Sigma f$ ، أى عدد المشاهدات أو إجمالى التكرارات . وفى التوزيعات الاحتمالية ، عادة مايشار إلى الوسط الحسابى μ بتعبير « القيمة المتوقعة » ، $E(X)$. ويمكن اشتقاق معادلة μ أو $E(X)$ للتوزيع الاحتمالى المنفصل بدءاً بمعادلة (٢ – ٢ أ) ووضع $f = P(X)$ ، وهى احتمال كل من النواتج الممكنة ، X . فيكون ، $\Sigma fX = \Sigma XP(X)$ ، والتى تمثل مجموع قيمة كل ناتج مضروبة فى احتمال حدوثها ، حيث $N = \Sigma f = \Sigma P(X)$ ، التى تمثل مجموع الاحتمالات لكل النواتج وتساوى 1 وعليه

$$E(X) = \mu = \sum XP(X) \qquad (٣ – ١٩)$$

(ب) معادلة التباين لبيانات المجتمع المبوبة (معادلة (٢ – ٩ أ)) هى

$$\sigma^2 = \frac{\sum f(X-\mu)^2}{N} \qquad (٣ – ٢٠)$$

مرة أخرى ، دع $f = P(X)$ احتمال كل النواتج ، $N = \Sigma f = \Sigma P(X) = 1$ ، فيمكن الحصول على معادلة التباين لتوزيع احتمالى منفصل :

$$\text{Var}\, X = \sigma_X^2 = \sum [X - E(X)]^2 P(X) = \sum X^2 P(X) - \left[\sum XP(X)\right]^2 = E(X^2) - [E(X)]^2 = E(X^2) - \mu^2$$

(٣ – ٢١)

٣ – ٢١ يعطى جدول (٣ – ٣) عدد طلبات الوظائف والتى عالجتها وكالة توظيف صغيرة خلال فترة 100 يوم الماضية . احسب العدد المتوقع للطلبات المعالجة وكذلك التباين والانحراف المعيارى .

جدول (٣ – ٣) عدد طلبات الوظائف المعدة خلال 100 يوم الماضية

عدد الأيام المتحققة فى المعالجة	عدد طالبى الوظـائف
10	7
10	8
20	10
30	11
20	12
10	14
100	

بقدر ما نعتبر أن خبرة المائة يوم الماضية نموذجية ، يمكننا إيجاد توزيع التكرار النسبى ، ومنه التوزيع الاحتمالى وتظهر هذه مع باقى الحسابات لإيجاد $E(X)$ و X فى جدول ؟ (٣ – ٤) :

طلبات مربعة $\operatorname{Var} X = \sigma_X^2 = \sum X^2P(X) - [XP(X)]^2 = 116 - (10.6)^2 = 116 - 112.36 = 3.64$

طلب $\operatorname{SD} X = \sigma_X = \sqrt{\sigma_X^2} = \sqrt{3.64} \cong 1.91$

جدول (٣ – ٤) الحسابات اللازمة لإيجاد القيمة المتوقعة والتباين

العدد X	الأيام f	$P(X)$	$XP(X)$	X^2	$X^2P(X)$
7	10	0.1	0.7	49	4.9
8	10	0.1	0.8	64	6.4
10	20	0.2	2.0	100	20.0
11	30	0.3	3.3	121	36.3
12	20	0.2	2.4	144	28.8
14	10	0.1	1.4	196	19.6
	$N = \sum f = 100$	$\sum P(X) = 1.0$	$\sum XP(X) = 10.6$		$\sum X^2P(X) = 116.0$
$E(X) = \mu = \sum XP(X) = 10.6$ طلب					

٣ – ٢٢ (أ) اذكر الشروط المطلوبة لتطبيق توزيع ذى الحدين (ب) ما احتمال 3 صور فى 5 رميات لعملة متوازنة ؟ (ج) ما احتمال أقل من 3 صور فى 5 رميات لعملة متوازنة ؟

(أ) يستخدم توزيع ذى الحدين لإيجاد احتمال $P(X)$ عدد X من النجاحات لحدث ما ، من بين عدد n من المحاولات لنفس التجربة عندما يكون (١) هناك فقط ناتجان متنافيان لكل محاولة (٢) المحاولات وعددها n مستقلة إحداها عن الأخرى ، (٣) احتمال النجاح p يبقى ثابتاً من محاولة إلى أخرى .

(ب)
$$P(X) = n\mathrm{C}_X p^X(1-p)^{n-X} = \binom{n}{X} p^X(1-p)^{n-X} = \frac{n!}{X!(n-X)!}p^X(1-p)^{n-X}$$

انظر المعادلات (٣ – ١٠) و (٣ – ١٧) . فى بعض الكتب تستخدم q بدلا من $1-p$ (احتمال الفشل) . هنا $n = 5$ ، $X = 3$ ، $p = 1/2$ و $1-p = 1/2$. بإحلال هذه القيم فى المعادلة السابقة ، نحصل على

$$P(3) = \frac{5!}{3!(5-3)!}(1/2)^3(1/2)^{5-3} = \frac{5!}{3!\,2!}(1/2)^3(1/2)^2 = \frac{5\cdot4\cdot3\cdot2\cdot1}{3\cdot2\cdot1\cdot2\cdot1}(1/2)^5 = 10(1/32) = 0.3125$$

(ج)

$$P(X < 3) = P(0) + P(1) + P(2)$$

$$P(0) = \frac{5!}{0!\,5!}(1/2)^0(1/2)^5 = \frac{1}{32} = 0.03125$$

$$P(1) = \frac{5!}{1!\,(5-1)!}(1/2)^1(1/2)^4 = \frac{5}{32} = 0.15625$$

$$P(2) = \frac{5!}{2!\,(5-2)!}(1/2)^2(1/2)^3 = \frac{10}{32} = 0.3125$$

أى ،

$$P(X < 3) = P(0) + P(1) + P(2) = 0.03125 + 0.15625 + 0.3125 = 0.5$$

٣ – ٢٣ (أ) افترض أن احتمال حصول أبوين على طفل أشقر الشعر هو $1/4$ فإذا كان فى الأسرة 6 أطفال ، ما احتمال أن نصفهم ذوو شعر أشقر ؟ (ب) إذا كان احتمال إصابة هدف بقذيفة واحدة هو 0.3 ، ماهو احتمال إصابة الهدف 3 مرات على الأقل خلال 4 قذائف ؟

(أ) هنا $n = 6$ ، $X = 3$ ، $p = 1/4$ و $1 - p = 3/4$. بإحلال هذه القيم فى صيغة ذى الحدين نحصل على:

$$P(3) = \frac{6!}{3!\,(6-3)!}(1/4)^3(3/4)^3 = \frac{6!}{3!\,3!}(1/64)(27/64) = \frac{6\cdot5\cdot4\cdot3\cdot2\cdot1}{3\cdot2\cdot1\cdot3\cdot2\cdot1}(27/4{,}096)$$

$$= 20\,\frac{27}{4{,}096} = \frac{540}{4{,}096} \cong 0.13$$

(ب) هنا $n = 4$ ، $X \geq 3$ ، $p = 0.3$ و $1 - p = 0.7$:

$$P(X \geqslant 3) = P(3) + P(4)$$

$$P(3) = \frac{4!}{3!\,(4-3)!}(0.3)^3(0.7)^1 = \frac{4\cdot3\cdot2\cdot1}{3\cdot2\cdot1\cdot1}(0.027)(0.7) = (4)(0.0189) = 0.0756$$

$$P(4) = \frac{4!}{4!\,(4-4)!}(0.3)^4(0.7)^0 = (0.3)^4 = 0.0081$$

وإذن

$$P(X \geqslant 3) = P(3) + P(4) = 0.0756 + 0.0081 = 0.0837$$

٣ – ٢٤ (أ) يأخذ مراقب جودة إنتاج عينة من 10 صمامات عشوائياً من شحنة كبيرة من الصمامات معروف أنها تحتوى على 20% صماماً معيباً . ماهو احتمال أن يكون عدد الصمامات المعيبة فى العينة أقل من أو يساوى 2 ؟

(ب) يأخذ مهندس فحص عينة من 15 وحدة عشوائياً من عملية إنتاج صناعى معروف أنها تنتج 85% وحدات مقبولة . ما احتمال أن تكون 10 من الوحدات المسحوبة مقبولة ؟

(أ) هنا $n = 10$ ، $X \leq 2$ ، $p = 0.2$ و $1 - p = 0.8$

$$P(X \leqslant 2) = P(0) + P(1) + P(2)$$

$P(0) = \frac{10!}{0!\,(10-0)!}(0.2)^0(0.8)^{10} = 0.1074$ (بالبحث عن $n = 10$ ، $X = 0$ و $p = 0.2$ فى ملحق ١)

$P(1) = 0.2684$ (بالبحث عن $n = 10$ ، $X = 1$ و $p = 0.2$ فى ملحق ١)

$P(2) = 0.3020$ (بالبحث عن $n = 10$ ، $X = 2$ ، $p = 0.2$ فى ملحق ١)

فيكون ،

$$P(X \leqslant 2) = P(0) + P(1) + P(2) = 0.1074 + 0.2684 + 0.3020 = 0.6778$$

(ب) هنا $n = 15$ ، $X = 10$ ، $p = 0.85$ و $1 - p = 0.15$. وحيث أن ملحق ١ يعطى احتمالات ذى الحدين حتى 0.5 فيجب تحوير هذه المسألة . فاحتمال $X = 10$ وحدات مقبولة مع $p = 0.85$ يعادل احتمال $X = 5$

وحدات معيبة عندما $p = 0.15$. باستخدام $n = 15$ ، $X = 5$ وحدات معيبة ، $p = 0.15$ نحصل على 0.0449 (من ملحق ١) .

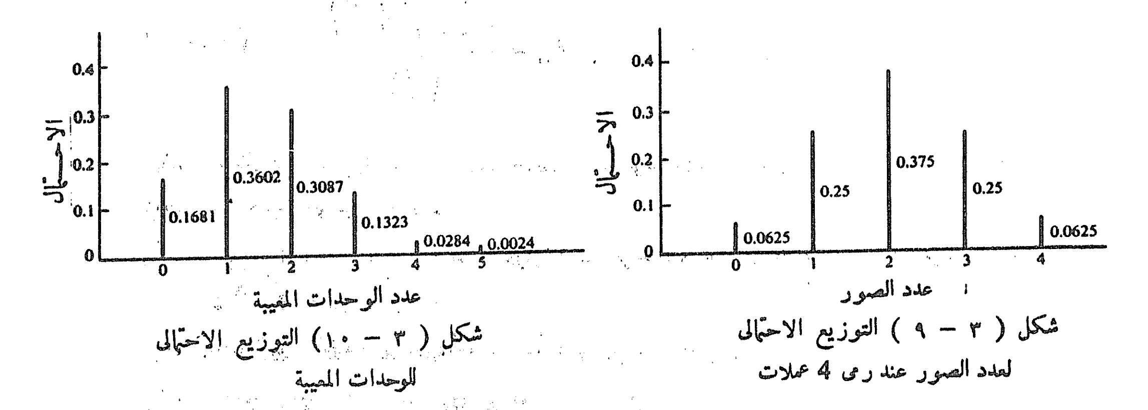
٣ – ٢٥ (أ) إذا ألقيت 4 عملات متوازنة فى آن واحد (أو ألقيت عملة واحدة 4 مرات) ، إحسب التوزيع الاحتمالى بأكمله وارسمه
(ب) احسب وارسم التوزيع الاحتمالى لعينة من 5 وحدات مسحوبة عشوائياً من عملية إنتاجية معروف أنها تنتج 30% وحدة معيبة .

(أ) باستخدام $n = 4$ ، $X = 0H, 1H, 2H, 3H$ أو $4H$ ، $p = 1/2$ وباستخدام ملحق ١ نحصل على

$$P(0H) = 0.0625, \quad P(1H) = 0.2500, \quad P(2H) = 0.3750, \quad P(3H) = 0.2500, \quad P(4H) = 0.0625$$

إذن ،
$$P(0H) + P(1H) + P(2H) + P(3H) + P(4H)$$
$$= 0.0625 + 0.2500 + 0.3750 + 0.2500 + 10.0625 = 1$$

انظر شكل (٣ – ٩) . لاحظ أن $p = 0.5$ وأن التوزيع الاحتمالى فى شكل (٣ – ٩) متماثل .



شكل (٣ – ١٠) التوزيع الاحتمالى للوحدات المعيبة

شكل (٣ – ٩) التوزيع الاحتمالى لعدد الصور عند رمى 4 عملات

(ب) باستخدام $n = 5$ ، $X = 0, 1, 2, 3, 4$ أو 5 وحدات معيبة ، و $p = 0.3$ ، نحصل على :

$$P(0) = 0.1681, \quad P(1) = 0.3602, \quad P(2) = 0.3087, \quad P(3) = 0.1323, \quad P(4) = 0.0284, \quad P(5) = 0.0024$$

إذن ،
$$P(0) + P(1) + P(2) + P(3) + P(4) + P(5)$$
$$= 0.1681 + 0.3602 + 0.3087 + 0.1323 + 0.0284 + 0.0024 = 1$$

انظر شكل (٣ – ١٠) . لاحظ أن $p < 0.5$ وأن التوزيع الاحتمالى فى شكل (٣ – ١٠) ملتو إلى اليمين .

٣ – ٢٦ احسب القيمة المتوقعة والانحراف المعيارى وحدد التماثل أو عدم التماثل للتوزيع الاحتمالى فى (أ) المسألة ٣ – ٢٣ (أ) (ب) المسألة ٣ – ٢٣ (ب) (جـ) المسألة ٣ – ٢٤ (أ) و (د) المسألة ٣ – ٢٤ (ب)

(أ)
$$E(X) = \mu = np = (6)(1/4) = 2/3 \cong 0.67 \text{ طفل أشقر}$$

$$\text{SD}\,X = \sqrt{n - p(1-p)} = \sqrt{6(1/4)(3/4)} = \sqrt{18/16} = \sqrt{1.125} \cong 1.06 \text{ طفل أشقر}$$

ولأن $p < 0.5$ فإن التوزيع الاحتمالى لعدد الأطفال الشقر ملتو إلى اليمين .

(ب)
$$E(X) = \mu = np = (4)(0.3) = 1.2 \text{ إصابة}$$

$$\text{SD}\,X = \sqrt{np(1-p)} = \sqrt{(4)(0.3)(0.7)} = \sqrt{0.84} \cong 0.92 \text{ إصابة}$$

لأن $p < 0.5$ ، التوزيع الاحتمالى ملتو لليمين .

(ج) صمامات معيبة $E(X) = \mu = np = (10)(0.2) = 2$

صمامات معيبة $\text{SD}\,X = \sqrt{np(1-p)} = \sqrt{(10)(0.2)(0.8)} = \sqrt{1.6} \cong 1.26$

لأن $p < 0.5$ ، التوزيع ملتو لليمين .

(د) وحدة مقبولة $E(X) = \mu = np = (15)(0.85) = 12.75$

وحدة مقبولة $\text{SD}\,X = \sqrt{np(1-p)} = \sqrt{15(0.85)(0.15)} = \sqrt{1.9125} \cong 1.38$

لأن $p > 0.5$ ، التوزيع الاحتمالى ملتو لليسار .

٣ – ٢٧ عندما تتم المعاينة من مجتمع محدود بدون إحلال ، لايمكن استخدام توزيع ذى الحدين لأن الأحداث تكون غير مستقلة . فى هذه الحالة تستخدم التوزيع فوق الهندسى . وقانونه كالآتى :

$$P_{\text{H}} = \frac{\binom{N - X_t}{n - X}\binom{X_t}{X}}{\binom{N}{n}} \quad \text{التوزيع فوق الهندسى} \qquad (٣-٢٢)$$

ويقيس احتمال عدد نجاحات X ، فى عينة حجمها n مأخوذة عشوائياً بدون إحلال من مجتمع حجمه N ، حيث X_t وحدة تتحقق فيها الخاصية التى سميت « نجاحاً »

(أ) باستخدام المعادلة ، احسب احتمال اختيار 2 من الرجال فى عينة من 6 أفراد اختيرت عشوائياً بدون إحلال من مجموعة من 10 أفراد منها 5 رجال .

(ب) كيف تكون النتيجة لو استخدمنا (خطأ) توزيع ذى الحدين ؟

(أ) هنا $X = 2$ ، $n = 6$ ، $N = 10$ ، $X_t = 5$:

$$P_{\text{H}} = \frac{\binom{10-5}{6-2}\binom{5}{2}}{\binom{10}{6}} = \frac{\binom{5}{4}\binom{5}{2}}{\binom{10}{6}} = \frac{\frac{5!}{4!\,1!}\,\frac{5!}{2!\,3!}}{\frac{10!}{6!\,4!}} = \frac{(5)(10)}{210} \cong 0.24$$

(ب)

$$P(2) = \frac{n!}{X!\,(n-X)!}\,p^X(1-p)^{n-X} = \frac{6!}{2!\,4!}(1/2)^2(1/2)^4 = \frac{15}{64} = 0.23$$

ويجب ملاحظة أنه عندما يكون حجم العينة صغيراً جداً بالنسبة لحجم المجتمع (مثلا ، أقل من 5% من المجتمع)، فإن المعاينة بدون إحلال يكون تأثيرها صغيراً على احتمال النجاح فى كل محاولة ويكون توزيع ذى الحدين (الأسهل فى الاستخدام) تقريباً جيداً للتوزيع فوق الهندسى . وهذا هو سبب استخدام توزيع ذى الحدين فى المسألة ٣–٢٤ (أ).

توزيع بواسون :

٣ – ٢٨ (أ) ما الفرق بين توزيع ذى الحدين وتوزيع بواسون ؟ (ب) متى يمكن تطبيق توزيع بواسون؟ اعط بعض الأمثلة على ذلك .

(ج) اذكر صيغة توزيع بواسون واشرح معنى الرموز المختلفة .

(د) تحت أى شروط يمكن استخدام توزيع بواسون كتقريب لتوزيع ذى الحدين ؟ لماذا يمكن أن يكون هذا مفيداً ؟

(أ) بينما يمكن استخدام توزيع ذى الحدين لإيجاد احتمال عدد معين من النجاحات من بين عدد n من المحاولات ، فإن توزيع بواسون يستخدم لإيجاد احتمال عدد معين من النجاحات فى وحدة الزمن . والشروط الأخرى لتطبيق توزيع ذى الحدين

مطلوبة أيضاً لتطبيق بواسون . وهذه هى (١) أن يكون هناك فقط ناتجان متنافيان . (٢) الأحداث يجب أن تكون مستقلة (٣) يبقى متوسط عدد مرات النجاح لوحدة الزمن ثابتاً .

(ب) يستخدم توزيع بواسون عادة فى بحوث العمليات لحل مشاكل الإدارة . وبعض الأمثلة هى عدد المكالمات التليفونية لمركز بوليس فى الساعة ، عدد العملاء الذين يصلون إلى طلمبة البنزين فى الساعة ، وعدد حوادث المرور عند تقاطع مافى الأسبوع .

(ج) احتمال عدد معين من النجاحات لوحدة الزمن $P(X)$ ، يمكن إيجاده باستخدام

$$P(X) = \frac{\lambda^X e^{-\lambda}}{X!}$$

حيث X = عدد معين من النجاحات

λ = (الحرف اليونانى لامدا) متوسط عدد النجاحات فى فترة زمنية معينة

e = أساس نظام اللوغاريتمات الطبيعية أو 0.271828

وبمعرفة قيمة λ يمكن إيجاد $e^{-\lambda}$ من ملحق ٢) ، وإحلالها فى المعادلة ، لإيجاد $P(X)$. لاحظ أن λ هى الوسط الحسابى والتباين لتوزيع بواسون .

(د) يمكن استخدام توزيع بواسون كتقريب لتوزيع ذى الحدين عندما عدد المحاولات n كبيراً ، p أو $1-p$ صغيرة (أحداث نادرة) . وكقاعدة عملية جيدة يستخدم توزيع بواسون عندما $n \geq 30$ و np أو $n(1-p)$ أقل من 5 ، إذ عندما تكون n كبيرة فإننا ستخدام توزيع ذى الحدين يستغرق وقتاً طويلا كما أن جداول احتمالات ذى الحدين قد لا تكون متاحة للقيم الصغيرة للاحتمال p .

٣ — ٢٩ تشير الخبرة السابقة إلى أنه فى المتوسط يتوقف 6 عملاء للتزود بالبنزين عند طلمبة بنزين كل ساعة . ماهو احتمال

(أ) توقف 3 عملاء فى ساعة ما ؟ (ب) 3 عملاء أو أقل فى ساعة ما ؟

(ج) ماهى القيمة المتوقعة ، أو الوسط الحسابى والانحراف المعيارى لهذا التوزيع ؟

(أ)

$$P(3) = \frac{6^3 e^{-6}}{3!} = \frac{(216)(0.00248)}{3 \cdot 2 \cdot 1} = \frac{0.53568}{6} = 0.08928$$

(ب)

$$P(X \leq 3) = P(0) + P(1) + P(2) + P(3)$$

$$P(0) = \frac{6^0 e^{-6}}{0!} = \frac{(1)(0.00248)}{1} = 0.00248$$

$$P(1) = \frac{6^1 e^{-6}}{1!} = \frac{(6)(0.00248)}{1} = 0.01488$$

$$P(2) = \frac{6^2 e^{-6}}{2!} = \frac{(36)(0.0248)}{2.1} = 0.04464$$

$$P(3) = 0.08928 \text{ (from part } a)$$

إذن ،

$$P(X \leq 3) = 0.00248 + 0.01488 + 0.04464 + 0.08928 = 0.15128$$

(ج) القيمة المتوقعة ، أو الوسط الحسابى ، لتوزيع بواسون هذا هو $\lambda = 6$ عملاء ، وانحراف معيارى $\sqrt{\lambda} = \sqrt{6} \cong 2.45$ عميل

٣ – ٣٠ توضح الخبرة الماضية أن 1% من مصابيح الكهرباء المنتجة فى مصنع ما هى مصابيح معيبة . فى عينة من 30 مصباح ، أوجد احتمال وجود أكثر من مصباح معيب باستخدام (أ) توزيع ذى الحدين (ب) توزيع بواسون كتقريب لذى الحدين .

(أ) عندما $n = 30$ ، $p = 0.01$ ومطلوب إيجاد $P(X > 1)$. باستخدام ملحق ١ نحصل على

$$P(2) + P(3) + P(4) + \cdots = 0.0328 + 0.0031 + 0.0002 = 0.0361, \text{ or } 3.61\%$$

(ب) حيث $n = 30$ و $np = (30)(0.01) = 0.3$ ، فيمكننا استخدام تقريب بواسون لذى الحدين . بوضع $\lambda = np = 0.3$ ، فيمكننا إيجاد $P(X > 1) = 1 - P(X \leq 1)$ ، حيث X عدد المصابيح المعيبة . باستخدام معادلة (٣ – ١٣) ، نحصل على

$$P(1) = \frac{0.3^1 e^{-0.3}}{1!} = (0.3)(0.74082) = 0.222246$$

$$P(0) = \frac{0.3^0 e^{-0.3}}{0!} = e^{-0.3} = 0.74082$$

$$P(X \leq 1) = P(1) + P(0) = 0.222246 + 0.74082 = 0.963066$$

إذن ،

$$P(X > 1) = 1 - P(X \leq 1) = 1 - 0.963066 = 0.036934, \text{ or } 3.69\%$$

ومع زيادة n فإن التقريب يقترب أكثر من الاحتمال باستخدام ذى الحدين .

التوزيعات الاحتمالية المتصلة : التوزيع الطبيعى :

٣ – ٣١ (أ) عرف ماذا يقصد بالمتغير المتصل واعط بعض الأمثلة . (ب) عرف ماذا يقصد بالتوزيع الاحتمالى المتصل . (جـ) اشتق صيغة القيمة المتوقعة والتباين لتوزيع احتمالى متصل .

(أ) المتغير المتصل هو الذى يمكن أن يأخذ أى قيمة داخل فترة معينة فيمكن ببساطة قياس المتغير المتصل بأى درجة من الدقة باستخدام وحدات قياس أصغر فأصغر . فمثلا ، يمكن إذا قلت أن عملية إنتاجية تأخذ 10 ساعات ، فهذا يعنى أى فترة بين 9.5 و10.4 ساعات (10 ساعات مقربة إلى أقرب ساعة) فإذا استخدمنا الدقيقة كوحدة قياس ، فيمكن القول أن العملية الإنتاجية تأخذ 10 ساعات و 20 دقيقة . وهذا يعنى أى فترت زمنية بين 10 ساعات ، 19.5 دقيقة و 10 ساعات ، 20.4 دقيقة وهكذا . ومن ثم يكون الزمن متغيراً متصلا ، ومثل ذلك ، الوزن ، المسافة ، والحرارة .

(ب) التوزيع الاحتمالى المتصل : يشير إلى مدى كل القيم الممكنة التى يمكن أن يأخذها متغير عشوائى متصل ، مع احتمالاتها المناظرة . وكثيراً مايسمى التوزيع الاحتمالى لمتغير عشوائى متصل بدالة كثافة الاحتمال ، أو ببساطة دالة الاحتمال . ويتم تمثيلها بمنحنى سلس بحيث أن مجموع المساحة (الاحتمال) تحت المنحنى يساوى 1 وحيث أن المتغير العشوائى المتصل يمكن أن يأخذ داخل فترة معينة عدداً لانهائياً من القيم ، فإن احتمال أن يأخذ المتغير المتصل قيمة معينة يكون 0 . ولكن ، يمكننا قياس احتمال أن يقع متغير عشوائى متصل X داخل فترة محددة (مثلا بين X_1 و X_2) باستخدام المساحة تحت المنحنى داخل هذه الفترة . أى ،

$$P(X_1 < X < X_2) = \int_{X_1}^{X_2} f(X)\,dX \qquad (٣ – ٢٢)$$

حيث $f(X)$ هى معادلة دالة كثافة الاحتمال ، وعلامة التكامل $\int$ تقابل علامة الجمع Σ للمتغيرات المنفصلة . ويوجد بالملاحق جداول احتمالات لبعض التوزيعات الاحتمالية المتصلة الأكثر استخداماً ، وعليه تنتفى الحاجة إلى إجراء التكامل بأنفسنا .

(جـ) يمكن اشتقاق القيمة المتوقعة أو الوسط الحسابى ، والتباين للتوزيعات الاحتمالية المتصلة بإحلال $\int$ محل Σ ، $f(X)$ محل $P(X)$ فى معادلات القيمة المتوقعة والتباين للتوزيعات الاحتمالية المنفصلة (معادلات) ٣ – ١٩ و ٣ – ٢٠) أى ،

$$E(X) = \mu = \int Xf(X)\, dX \qquad (٣ - ٢٤)$$

$$\text{Var}\, X = \sigma^2 = \int [X - E(X)]^2 f(X)\, dX \qquad (٣ - ٢٥)$$

٣ - ٣٢ (أ) ماهو التوزيع الطبيعي ؟ (ب) ما استخداماته ؟ (ـ) ماهو التوزيع الطبيعي القياسي ؟ ما استخداماته ؟

(أ) التوزيع الطبيعي هو دالة احتمال متصلة وهو جرسى الشكل ، متماثل حول الوسط الحسابي ، ومعتدل (أنظر التعريف فى قسم ٢ - ٤) . وكلما تحركنا بعيداً عن الوسط الحسابي فى كلا الاتجاهين ، اقترب منحنى التوزيع الطبيعي من المحور الأفقى (ولكنه لايلمسه أبداً) . ومعادلة دالة الاحتمال الطبيعي هى

$$f(X) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2 \right] \qquad (٣ - ٢٦)$$

حيث $f(X)$ = ارتفاع المنحنى الطبيعي
exp = الثابت 2.7183
π = الثابت 3.1416
μ = الوسط الحسابي للتوزيع
σ = الانحراف المعيارى للتوزيع

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2 \right] dX = 1$$

(المساحة الكلية تحت المنحنى الطبيعي من ناقص ما لانهاية إلى زائد مالا نهاية)

(ب) إن التوزيع الطبيعي هو أكثر التوزيعات الاحتمالية استخداماً فى التحليل الإحصائى . فكثير من التوزيعات الموجودة فعلا فى الطبيعة وفى الصناعة طبيعية . وأمثلة ذلك ، مقاييس الذكاء ، الأوزان ، والأطوال لعدد كبير من الناس ، والتفاوتات فى أبعاد عدد كبير من الأجزاء التى تنتجها ماكينة ما . كما يستخدم التوزيع الطبيعي كثيراً كتقريب لتوزيعات أخرى مثل توزيع ذى الحدين وتوزيع بواسون (أنظر المسائل ٣ - ٣٧ و ٣ - ٣٨) . وعادة ما يكون توزيع متوسط العينات ، والنسب طبيعياً ، بصرف النظر عن شكل توزيع المجتمعات الأصلية (أنظر قسم ٤ - ٢) .

(جـ) التوزيع الطبيعي القياسى هو توزيع طبيعي فيه $\mu = 0$ و $\sigma^2 = 1$. ويمكن تحويل أى توزيع طبيعي (معرفاً بقيم معينة μ و σ^2) إلى توزيع طبيعي معيارى بوضع $\mu = 0$ ونقيس الانحرافات عن μ بوحدات من الانحراف المعيارى . وعادة يمكن إيجاد المساحات (الاحتمالات) بتحويل قيم X إلى قيم z المناظرة (أى أن ، $z = (X - \mu)/\sigma$) والبحث عن قيم z هذه فى ملحق (٣) .

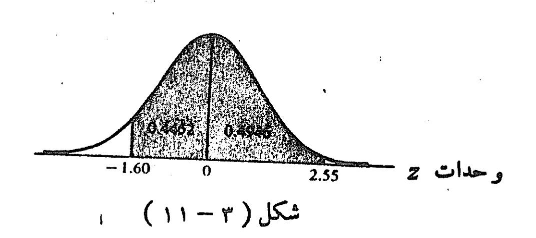
٣ - ٣٣ أوجد المساحة تحت التوزيع الطبيعي القياسى (أ) بين $z \pm 1$ ، $z \pm 2$ و $z \pm 3$ ، (ب) من $z = 0$ إلى $z = 0.88$ ، (جـ) من $z = 1.60$ إلى $z = 2.55$ ، (د) إلى اليسار من $z = 1.60$ ، (هـ) إلى اليمين من $z = 2.55$ ، (و) إلى اليسار من $z = -1.60$ وإلى اليمين من $z = 2.55$.

(أ) المساحة (الاحتمال) الداخلة تحت المنحنى الطبيعي القياسى بين $z = 0$ و $z = 1$ يتم الحصول عليها بالبحث عن القيمة 1.0 فى ملحق (٣) . ويتم هذا بالتحرك إلى أسفل فى عمود z فى الجدول حتى نصل إلى صف الرقم 1.0 فنتحرك فيه حتى نصل إلى العمود الذى عنوانه 00 . القيمة التى نحصل عليها هى 0.3413 . وهذا يعنى أن 34.13% من المساحة الكلية (من 1 ، أو 100%) تحت المنحنى تقع بين $z = 0$ و $z = 1.00$. وكنتيجة التماثل ، فإن المساحة بين $z = 0$ و $z = -1$ هى أيضاً 0.3413 ، أو 34.13% . ومن ثم فإن المساحة بين $z = -1$ و $z = 1$ هى 68.26% (أنظر شكل ٣ - ٤) . وبالمثل ، فان المساحة بين $z = 0$ و $z = 2$ هى 0.4772 أو 47.72% (بالبحث عن

قيمة $z = 2.00$ فى الجدول) ، فتكون المساحة بين $z = \pm 2$ هى 95.54 % (أنظر شكل ٢ – ٤) . والمساحة بين $z = \pm 3$ 99.74% (أنظر شكل ٣ – ٤) . لاحظ أن الجدول يعطى قيما تفصيلية عن z فقط حتى 2.99 لأن المساحة تحت المنحنى خارج $z = \pm 3$ صغيرة جداً ويمكن إهمالها .

(ب) المساحة بين $z = 0$ و $z = 0.88$ يمكن الحصول عليها بالبحث بالجدول مقابل 0.88 . وهذا يعطى القيمة 0.3106 .

(جـ) المساحة بين $z = 0$ و $z = -1.60$ يمكن إيجادها بالبحث بالجدول مقابل $z = 1.60$. الرقم المقابل 0.4452 . المساحة بين $z = 0$ و $z = 2.55$ يمكن إيجادها بالجدول مقابل $z = 2.55$. الرقم المقابل 0.4946 . ومن ثم فإن المساحة بين $z = -1.60$ و $z = 2.55$ تساوى 0.4452 زائداً 0.4946 أى 0.9398 أو 93.98% (أنظر شكل ٣ – ١١) . وفى كل المسائل من هذا النوع يكون من المفيد الاستعانة برسم للتوزيع .



شكل (٣ – ١١)

(د) نحن نعلم أن اجمالى المساحة تحت المنحنى الطبيعى القياسى تساوى 1 . وكنتيجة للتماثل فإن 0.5 من المساحة تقع على كل من جانبى $\mu = 0$ وحيث أن 0.4452 تمتد بين $z = 0$ و $z = -1.60$ ، فان $0.5 - 0.4452 = 0.0548$ أو 5.48% ، تكون هى المساحة فى الطرف الأيسر ، إلى اليسار من -1.60 (أنظر شكل ٣ – ١١) .

(هـ) $0.5 - 0.4946 = 0.0054$ أو 0.54 % هى المساحة فى الطرف الأيمن ، إلى اليمين من $z = 2.55$ (أنظر شكل ٣ – ١١)

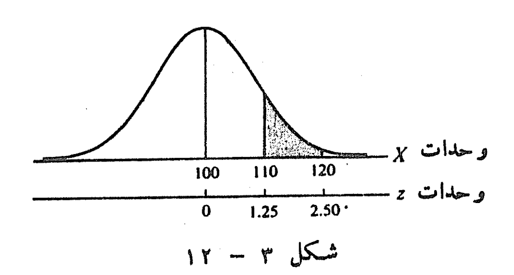
(و) المساحة إلى اليسار من $z = -1.60$ وإلى اليمين من $z = 2.455$ تساوى $1 - 0.9398$ (من إجابة جـ) . أى 0.0602 أو 6.02 % من الإجمالى .

٣ – ٣٤ إذا كان من المعروف أن توزيع عمر المصابيح الكهربائية يتبع التوزيع الطبيعى حيث $\mu = 100$ h و $\sigma = 8$ h . ما احتمال أن مصباحاً اختير عشوائياً له عمر بين 110 و 120 ساعات احتراق ؟

المطلوب هنا إيجاد $P(110 < X < 120)$ ، حيث تشير X إلى زمن الإضاءة مقيساً بالساعات . بمعلومية $\mu = 100$ h و $\sigma = 8$h ، وبوضع $X_1 = 110$ h و $X_2 = 120$ h ، نحصل على

$$z_1 = \frac{X_1 - \mu}{\sigma} = \frac{110 - 100}{8} = 1.25 \quad \text{و} \quad z_2 = \frac{X_2 - \mu}{\sigma} = \frac{120 - 100}{8} = 2.50$$

أى أننا نرغب فى إيجاد المساحة (الاحتمال) بين $z_1 = 1.25$ و $z_2 = 2.50$ (المساحة المظللة فى شكل ٣ – ١٢) . بالبحث مقابل $z_2 = 2.50$ فى ملحق (٣) ، فنحصل على 0.4938 . وهذه هى المساحة بين $z = 0$ و $z_2 = 2.50$. وبالبحث مقابل $z_1 = 1.25$ ، نحصل على 0.3944 . وهذه هى المساحة بين $z = 0$ و $z = 1.25$. بطرح 0.3944 من 0.4938 ، نحصل على 0.0994 أو 9.94% ، للمساحة المظللة التى تعطى $P(110 < X < 120)$.



شكل ٣ - ١٢

٣ - ٣٥ افترض أن دخول الأسر تتبع التوزيع الطبيعى مع $\mu = \$ 16,000$ و $\sigma = \$ 2,000$. ما احتمال أن يكون دخل أسرة اختيرت عشوائياً (أ) بين $ 15,000 و $ 18,000 ؟ (ب) أقل من $ 15,000 ؟ (جـ) أعلى من $ 18,000 ؟ (د) أعلى من $ 20,000 ؟

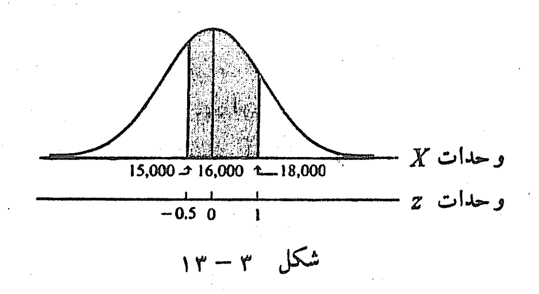
(أ) نحن نرغب فى إيجاد $P(\$ 15,000 < X < \$ 18,000)$ حيث X دخل الأسرة :

$$z_1 = \frac{X_1 - \mu}{\sigma} = \frac{\$15,000 - \$16,000}{\$2,000} = -0.5 \quad \text{و} \quad z_2 = \frac{X_2 - \mu}{\sigma} = \frac{\$18,000 - \$16,000}{\$2,000} = 1$$

أى أننا نرغب فى إيجاد المساحة (الاحتمال) بين $z_1 = -0.5$ و $z_2 = 1$ (المساحة المظللة فى شكل ٣ - ١٣) .
بالبحث مقابل $z = 0.5$ فى ملحق ٣ ، نحصل على 0.1915 للمساحة بين $z = 0$ و $z = -0.5$. بالبحث مقابل $z = 1$ ، نحصل على 0.3413 للمساحة بين $z = 0$ و $z = 1$. ومن ثم فان
$P(\$ 15,000 \leq X \leq \$ 18,000) = 0 . 1915 + 0.3413 = 0.5328$ أو 53.28% .

(ب) $P(X < \$ 15,000) = 0.5 - 0.151 = 0.3085$ أو 30.85% (المساحة غير المظللة فى الطرف الأيسر فى شكل ٣ - ١٣) .

(جـ) $P(X < \$ 15,000) = 0. 5 - 0.1915 = 0.3085$ أو 15.87% (المساحة غير المظللة فى الطرف الأيمن فى شكل ٣ - ١٣) .



شكل ٣ - ١٣

(د) $X = \$ 20,000$ تناظر $z = (\$ 20,000 - 16,000)/\$2,000 = 2$ ومن ثم
$P(X > \$ 20,000) = 0.5 - 0 4772 = 0.0228$ أو 2.28% .

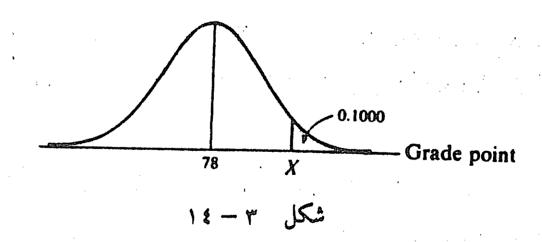
٣ - ٣٦ درجات امتحان نصف العام فى مادة الإحصاء لفصل كبير موزعة طبيعياً بوسط حسابى 78 وانحراف معيارى 8 ، ويريد الأستاذ أن يعطى تقدير A لنسبة 10% من الطلاب . ماهو الحد الأدنى للدرجات الذى يعطى تقدير A فى امتحان نصف العام ؟

فى هذه المسألة المطلوب تحديد الدرجة التى يزيد عنها 10% من الطلاب . وهذا يتضمن تحديد النقطة X بحيث أن 10% من المساحة تحت المنحنى الطبيعى تقع إلى اليمين من X (المساحة المظللة فى شكل ٣ - ١٤) . وحيث أن إجمالى المساحة تحت

المنحنى إلى اليمين من 78 هى 0.5 فإن المساحة غير المظللة فى شكل ٣ - ١٤ (إلى اليمين من 78 تكون 0.4 نحتاج إذن أن نبحث داخل الجدول داخل الجدول فى ملحق (٣) عن أقرب قيمة إلى 0.4 . وهذه هى 0.3997 والتى تقابل z تساوى 1.28 . وقيمة الدرجة المناظرة لقيمة z مساوية 1.28 يمكن إيجادها بإحلال القيم المعروفة فى $z = (X - \mu)/\sigma$ والحل لإيجاد X :

$$1.28 = \frac{X - 78}{8}$$

وهذا يعطى $X - 78 = 10.24$ ومن ثم فإن $X = 78 + 10.24 = 88.24$ أو 88 لأقرب درجة صحيحة .



شكل ٣ - ١٤

٣ - ٣٧ تشير الخبرة الماضية أن 30% من الناس الذين يدخلون محلا ما يقومون بالشراء فعلا . باستخدام (أ) توزيع ذى الحدين (ب) التوزيع الطبيعى كتقريب لذى الحدين ، أوجد احتمال أنه من بين 30 شخصاً يدخلون المحل ، فإن 10 أشخاص أو أكثر سوف يقومون بالشراء فعلا .

(أ) هنا $n = 30$ ، $p = 0.3$ ، $1 - p = 0.7$ والمطلوب إيجاد $P(X \geq 10)$. باستخدام ملحق (١) (جدول احتمالات ذى الحدين) ،

$$(X \geq 10) = P(10) + P(11) + P(12) + \cdots + P(30) = 0.1416 + 0.1103 + 0.0749 + 0.0444 + 0.0231$$
$$+ 0.0106 + 0.0042 + 0.0015 + 0.005 + 0.001$$
$$= 0.4112$$

(ب) أشخاص $\mu = np = (30)(0.3) = 9$ persons, and $\sigma = \sqrt{np(1-p)} = \sqrt{(30)(0.3)(0.7)} = \sqrt{6.3} \cong 2.51$

وحيث أن $n = 30$ وكلا من mp و $n(1 - p)$ أكبر من 5 فيمكن تقريب احتمالات ذى الحدين باستخدام احتمالات التوزيع الطبيعى . ولكن ، عدد الأشخاص متغير منفصل . فلكى نستطيع استخدام التوزيع الطبيعى ، فيمكن التعامل مع عدد الأشخاص كما لو كان متغيراً متصلا وإيجاد $P(X \geq 9.5)$. فيكون ،

$$z = \frac{X - \mu}{\sigma} = \frac{9.5 - 9}{2.51} = \frac{0.5}{2.51} \cong 0.20$$

وبالبحث مقابل $z = 0.20$ ، نحصل على 0.0793 (من ملحق (٣)) . وهذا يعنى أن 0.0793 من المساحة تحت المنحنى الطبيعى تقع بين $z = 0$ و $z = 0.20$. فيكون $P(X \geq 9.5) = 0.5 - 0.0793 = 0.4207$ (التقريب باستخدام التوزيع الطبيعى) . ومع كبر حجم n يتحسن التقريب المستخدم . (لو لم نعامل عدد الأشخاص كتغير متصل لوجدنا أن $P(X \geq 10) = 0.34$ ، ولا يكون التقريب الناتج على نفس الدرجة من الدقة) .

٣ - ٣٨ تنتج عملية إنتاجية 10 وحدات معيبة فى الساعة . أوجد احتمال أن 4 وحدات أو أقل تكون معيبة من بين إنتاج ساعة مختارة عشوائياً باستخدام (أ) توزيع بواسون (ب) التوزيع الطبيعى كتقريب لبواسون .

(أ) هنا $\lambda = 10$ والمطلوب إيجاد $P(X \leq 4)$ ، حيث X هى عدد الوحدات المعيبة من إنتاج ساعة مختارة عشوائياً . قيمة e^{-10} من ملحق (٢) هى 0.00005 . فيكون ،

$$P(0) = \frac{\lambda^0 e^{-10}}{0!} = e^{-10} = 0.00005$$

$$P(1) = \frac{\lambda^1 e^{-10}}{1!} = \frac{10(0.00005)}{1} = 0.0005$$

$$P(2) = \frac{\lambda^2 e^{-10}}{2!} = \frac{10^2(0.00005)}{2} = 0.0025$$

$$P(3) = \frac{\lambda^3 e^{-10}}{3!} = \frac{10^3(0.00005)}{6} = 0.0083335$$

$$P(4) = \frac{\lambda^4 e^{-10}}{4!} = \frac{10^4(0.00005)}{24} = 0.0208335$$

$P(X \leqslant 4) = P(0) + P(1) + P(2) + P(3) + P(4) = 0.00005 + 0.0005 + 0.0025 + 0.0083335 + 0.0208335$
$= 0.032217$, أو حوالى 3.23%

(ب) معاملة الوحدات كمتغير متصل (أنظر المسألة ٣ ــ ٣٧ (ب)) ، المطلوب هو إيجاد $P(X \leq 4.5)$ ، حيث X عدد الوحدات المعيبة ، $\mu = \lambda = 10$ و $\sigma = \sqrt{\lambda} = \sqrt{10} \cong 3.16$. فيكون ،

$$z = \frac{X - \mu}{\sigma} = \frac{4.5 - 10}{3.16} = \frac{-5.5}{3.16} = -1.74$$

بالنسبة لمقابل $z = 1.74$ فى ملحق (٣) . نحصل على 0.4591 وهذا يعنى أن $0.5 - 0.5491 = 0.0409$ من المساحة (الاحتمال (تحت المنحنى الطبيعى القياسى تقع إلى اليسار من $z = - 1.74$ وعليه فإن $P(X \leq 4.5) =$ 0.0409 أو 4.09% . ومع تزايد قيمة λ ، يتحسن التقريب المستخدم . (لو لم نعامل عدد الوحدات المعيبة كمتغير متصل ، لوجدنا أن $P(X \leq 4) = 0.0287$.

٣ ــ ٣٩ إذا كانت الأحداث أو النجاحات تتبع توزيع بواسون ، فيمكننا إيجاد احتمال أن يقع الحدث الأول خلال فترة زمنية معينة ، $P(T \leqq t)$ ، باستخدام التوزيع الاحتمالى الأسى . ولأننا نتعامل مع الزمن ، فإن التوزيع الأسى هو توزيع احتمال متصل . ويعبر عنه بالآتى :

$$P(T \leqslant t) = 1 - e^{-\lambda} \qquad (٣ - ٢٧)$$

حيث λ هو متوسط عدد مرات الحدث للفترة المعينة ، ويمكن الحصول على $e^{-\lambda}$ من ملحق (٢) . وتكون القيمة المتوقعة والتباين :

$$E(T) = \frac{1}{\lambda} \qquad (٣ - ٢٨)$$

$$\text{Var}\, T = \frac{1}{\lambda^2} \qquad (٣ - ٢٩)$$

(أ) فى تقرير المسألة (٣ ــ ٢٩) ، أوجد احتمال أنه بدءاً من نقطة زمنية عشوائية فإن العميل الأول سوف يتوقف عند طلمبة البنزين خلال نصف ساعة .

(ب) احتمال أنه لن يتوقف عميل عند طلمبة البنزين خلال نصف ساعة ؟

(جـ) ما القيمة المتوقعة ، والتباين للتوزيع الأسى ، حيث المتغير المستمر هو الزمن T

(أ) حيث أنه فى المتوسط يتوقف 6 عملاء عند طلمبة البنزين كل ساعة ، فإن $\lambda = 3$ متوسط عدد العملاء كل نصف ساعة . احتمال أن العميل الأول سوف يتوقف خلال نصف الساعة الأولى هو :

(أ) من ملحق (٣) $1 - e^{-\lambda} = 1 - e^{-3} = 1 - 0.04979$ (from App. 2) $= 0.9502$, or 95.02%

(ب) احتمال ألا يتوقف عميل خلال نصف ساعة

$$e^{-\lambda} = e^{-3} = 0.04979$$

(جـ) $E(T) = 1/\lambda = 1/6 \cong 0.17$ h للسيارة $T = 1/\lambda^2 = 1/36 \cong 0.03$ h للسيارة المربعة . ويمكن استخدام التوزيع الأسى أيضاً لحساب الوقت الذى يمر بين كل حدثين متتاليين .

٣ – ٤٠ الوسط الحسابى لسنوات التعليم لمجتمع ماهو 8 سنوات والانحراف المعيارى سنة واحدة . ما احتمال أن شخصاً ما تم اختياره عشوائياً من المجتمع تقع سنوات تعليمه بين 6 و 10 سنوات ؟ أقل من 6 سنوات أو أكثر من 10 سنوات ؟

حيث أننا لم نعط بيانات عن شكل التوزيع ، فيمكننا استخدام نظرية تشبتشيف والتى تنطبق على أى توزيع . باستخدام $\mu = 8$ و $\sigma = 1$ ، فإن 6 سنوات تقل عن الوسط الحسابى بعدد 2 انحراف معيارى . و 10 سنوات سنوات تزيد عن المتوسط بعدد 2 انحراف معيارى . باستخدام نظرية أو متباينة تشبتشيف :

$$P(|\bar{X} - \mu| < K\sigma) \geq 1 - \frac{1}{K^2} \qquad (٣٠ - ٣)$$

احتمال أن شخصاً ما تم اختياره عشوائياً من المجتمع تقع سنوات تعليمه بين 2 انحراف معيارى من المتوسط فى كل اتجاه هو :

$$1 - \frac{1}{K^2} = 1 - \frac{1}{2^2} = \frac{3}{4}, \text{ or } 75\%$$

وبالتالى ، فإن احتمال أن هذا الشخص تكون سنوات تعليمه أقل من 6 سنوات أو أكثر من 10 سنوات هو 25% .

مسائل إضافية

احتمال حدث منفرد :

٣ – ٤١ أى مدخل للاحتمالات تتضمنه العبارات الآتية ؟

(أ) احتمال صورة فى رمية لعملة متوازنة هو 1/2

(ب) التكرار النسبى للصورة فى 100 رمية لعملة هو 0.53

(جـ) احتمال أن تمطر غداً هو 20% .

الإجابة (أ) المدخل الكلاسيكى أو المسبق (ب) مدخل التكرار النسبى أو المدخل التجريبى (جـ) المدخل الشخصى أو الذاتى .

٣ – ٤٢ عند رمى عملة متوازنة ما احتمال (أ) كتابة (ب) صورة (جـ) ليس كتابة أو (د) كتابة أو ليس كتابة .

الإجابة : (أ) $P(T) = ½$ (ب) $P(H) = ½$ (جـ) $P(T') = 1/2$ أو (د) $P(T) + P(T') = 1$.

٣ – ٤٣ ما احتمال فى رمية واحدة لنردة أن نحصل على (أ) 1 (ب) 6 (جـ) ليس 1 أو (د) 1 أو ليس 1 ؟

الإجابة : (أ) $P(1) = 1/6$ (ب) $P(6) = 1/6$ (جـ) $P(1') = 5/6$ (د) $P(1) + P(1') = 1$.

٣ – ٤٤ عند سحب ورقة واحدة من مجموعة أوراق لعب ما احتمال أن نسحب (أ) كوبة (ب) آس (جـ) آس كوبة (د) ليس كوبة أو (هـ) كوبة أو ليس كوبة .

الإجابة : (أ) $P(C)\ 13/52 = 1/4$ (ب) $P(A)\ 4/52 = 1/13$ (جـ) $P(A_C) = 1/52$ (د) $P(C')\ 3/4$ (هـ) $P(C) + P(C') = 1$.

٣ — ٤٥ يحتوى وعاء على 12 كرة متماثلة تماماً باستثناء أن 4 منها زرقاء (B) ، 3 منها حمراء (R) ، 3 منها خضراء (G) ، و 2 منها بيضاء (W). عند سحب كرة واحدة ما احتمال أن تكون الكرة (أ) زرقاء ؟ (ب) حمراء ؟ (جـ) خضراء ؟ (د) بيضاء ؟ (هـ) ليست حمراء ؟ (و) ليست بيضاء ؟ (ز) بيضاء أو ليست بيضاء ؟ (ح) ماهى معاملات الترجيح لصالح سحب كرة خضراء ؟ (ط) ماهى معاملات الترجيح لسحب كرة ليست خضراء ؟

الإجابة (أ) $P(B) = 1/3$ أو 0.33 (ب) $P(R) = 1/4$ أو 0.25 (جـ) $P(G) = 1/4$ أو 0.25
(د) $P(W) = 1/6$ أو 0.167 (هـ) $P(R') = 0.75$ (و) $P(W') = 0.833$ (ز) $P(W) + P(W') = 1$
(ح) 3 : 9 (ط) 3 : 9

٣ — ٤٦ افترض أننا سحبنا ورقة من مجموعة لعب عادية ، ثم أعيدت الورقة إلى المجموعة وأعيد خلطها جيداً وسحبت ورقة ثانية . عندما أعيدت هذه العملية 520 مرة حصلنا على 136 بستونى . (أ) ماهو التكرار النسبى أو الاحتمال التجريبى للحصول على بستونى ؟ (ب) ماهو الاحتمال الكلاسيكى أو المسبق للحصول على بستونى ؟ (جـ) ماذا تتوقع أن يكون عليه التكرار النسبى ، أو الاحتمال التجريبى للحصول على بستونى فيما لو أعيدت العملية لمرات أخرى كثيرة ؟

الإجابة (أ) $136/520 \cong 0.26$ (ب) $P(S) = 1/4$ (جـ) أن يقترب من 1/4 أو 0.25 .

٣ — ٤٧ وجدت شركة تأمين أنه من بين عينة من 10,000 رجل بين سن 30 و سن 40 ، أصيب 87 رجلا بمرض خطير خلال سنة واحدة (أ) ماهو التكرار النسبى أو الاحتمال التجريبى لأن يصاب شخص بين سن 30 وسن 40 بمرض خطير خلال عام واحد ؟ (ب) لماذا تهتم شركات التأمين بهذه النتائج ؟ (جـ) افترض أن الشركة قد باعت تأميناً صحياً لعدد 1,387,684 رجلا بين سن 30 وسن 40 ، كم مطالبة تتوقعها الشركة خلال فترة العام ؟

الإجابة (أ) التكرار النسبى و الاحتمال التجريبى $87/10{,}000 = 0.0087$
(ب) تهتم شركة التأمين بالتكرار النسبى أو الاحتمال التجريبى لتحديد قسط التأمين .
(جـ) 12,073 لأقرب شخص .

احتمال الأحداث المتعددة :

٣ — ٤٨ ماهى نواع الأحداث الآتية ؟ (أ) سحب دينارى أو كوبة عند سحب ورقة واحدة من مجموعةأو راق لعب (ب) سحب دينارى أو بنتعند سحب ورقة واحدة من المجموعة (جـ) رميتان متتاليتان لعملة متوازنة (د) رميتان متتاليتان لنردة (هـ) سحب ورقتين من مجموعة أوراق لعب مع الإحلال (و) سحب ورقتين من المجموعة بدون احلال (ز) سحب كرتين من وعاء بدون إحلال .

الإجابة (أ) متنافية (ب) غير متنافية (جـ) مستقلة (د) مستقلة (هـ) مستقلة (و) غير مستقلة (ز) غير مستقلة .

٣ — ٤٩ ما احتمال الحصول على :
(أ) أربعة أو كثر عند رمى نردة مرة واحدة ؟ (ب) آس أو شايب عند سحب ورقة واحدة من مجموعة أوراق لعب
(جـ) كرة خضراء أو بيضاء من الوعاء في المسألة (٣ — ٤٥) ؟
الإجابة (أ) 1/2 (ب) 8/52 و 2/13 (جـ) 5/12

٣ — ٥٠ ما هو احتمال الحصول على :
(أ) دينارى و بنت عند سحب ورقة واحدة من مجموعة أوراق اللعب . (ب) دينارى أو بنت أو شايب ؟
(جـ) إنسان أسود أو امرأة كرئيس للولايات المتحدة إذا كان احتمال رئيس أسود 0.25 ، واحتمال امرأة 0.15 ، واحتمال امرأة سوداء 0.07 ؟

الإجابة (أ) 16/52 أو 4/13 (ب) 19/52 (جـ) 0.33

٣ — ٥١ ما احتمال الحصول على :

(أ) واحد وواحد فى رميتين لنردة ؟
(ب) ثلاث كتابات فى 3 رميات للعملة ؟
(ج) مجموع 6 عند رمى نردتين معاً ؟
(د) مجموع أقل من 5 عند رمى نردتين معا ؟
(هـ) مجموع 10 أو أكثر عند رمى نردتين معاً ؟

الإجابة (أ) 1/36 (ب) 1/8 (ج) 5/36 (د) 1/6 (هـ) 1/6

٣ — ٥٢ ما احتمال الحصول من مجموعة أوراق لعب على :

(أ) دينارى عند سحب الورقة الثانية علماً بأن الورقة الأولى التى سحبت ولم تعد كانت دينارى ؟
(ب) دينارى عند سحب الورقة الثانية علماً بأن الورقة الأولى التى سحبت ولم تعد لم تكن دينارى ؟
(ج) شايب عند سحب الورقة الثالثة علماً بأن الورقتين الأولى والثانية لم تعادا بعد السحب وكانتا بنتاً وولداً ؟ .

الإجابة (أ) 12/51 (ب) 13/51 (ج) 4/50

٣ — ٥٣ ما احتمال التقاط :

(أ) شايب كوبة ودينارى على الترتيب عند سحب ورقتين من مجموعة بدون إحلال ؟
(ب) كرة بيضاء وكرة خضراء على الترتيب عند سحب كرتين بدون إحلال من الوعاء فى المسألة (٣ — ٤٥) ؟
(ج) كرة خضراء وكرة بيضاء على الترتيب عند سحب كرتين بدون إحلال من الوعاء فى المسألة (٣ — ٤٥) ؟
(د) كرة خضراء وكرة بيضاء بأى ترتيب عند سحب كرتين بدون إحلال من نفس الوعاء ؟
(هـ) ثلاث كرات خضراء عند سحب ثلاث كرات بدون إحلال من الوعاء ؟

الإجابة (أ) 13/2,652 أو 1/204 (ب) 6/132 أو 1/22 (ج) 1/22 (د) 1/11 (هـ) 6/1320 أو 1/220

٣ — ٥٤ افترض أن احتمال أن تمطر السماء فى يوم معين هو 0.1 ، واحتمال أن يحدث لى حادث سيارة فى أى يوم هو 0.005 و 0.012 فى اليوم المطير .

(أ) ماهى القاعدة التى يجب أن أستخدمها لحساب احتمال أنه فى يوم معين سوف تمطر وأنه سوف يحدث لى حادث سيارة ؟
(ب) اكتب القاعدة المطلوبة فى (أ) باستخدام A للدلالة على حادثة ، R للدلالة على مطر .
(ج) احسب الاحتمال المطلوب فى (أ) .

الإجابة (أ) قاعدة الضرب للأحداث غير المستقلة (ب) $P(\text{R و A}) = P(\text{R}) . P(\text{A/R})$ (ج) 0.0012

٣ — ٥٥ (أ) ماهى القاعدة أو النظرية التى يجب استخدامها لحساب احتمال أنها كانت تمطر عندما حدثت لى حادثة سيارة فى المسألة (٣ — ٥٤) ؟

(ب) اذكر القاعدة أو النظرية التى تنطبق على (أ) . (ج) أجب على السؤال فى (ج) .

الإجابة (أ) نظرية بيز (ب) $P(\text{R/A}) = P(\text{R}). P(\text{A/R})/P(\text{A})$ (ج) 0.24

٣ — ٥٦ كم عدد الطرق التى يمكن بها تخصيص 6 أشخاص مؤهلين فى (أ) ثلاث وظائف تدريبية متاحة إذا كانت الوظائف متماثلة تماماً ؟
(ب) ثلاث وظائف تدريبية مختلفة ؟ (ج) ست وظائف تدريبية مختلفة .

الإجابة (أ) 20 (ب) 120 (ج) 720

التوزيعات الاحتمالية المنفصلة : توزيع ذى الحدين :

٣ – ٥٧ يظهر التوزيع الاحتمالى لمرتادى أحد المطاعم للغذاء فى جدول (٣ – ٥) . احسب
(أ) العدد المتوقع للمترددين على المطعم وقت الغذاء (ب) التباين (جـ) الانحراف المعيارى .

جدول (٣ – ٥) التوزيع الاحتمالى للمترددين على أحد المطاعم وقت الغذاء

عدد المترددين ، X	الاحتمال $P(X)$
100	0.2
110	0.3
118	0.2
120	0.2
125	0.1
	1.0

الإجابة (أ) 113.1 متردد (ب) 65.69 متردد مربع (جـ) 8.10 متردد

٣ – ٥٨ ما احتمال :
(أ) الحصول على 4 صورة ، 2 كتابة فى 6 رميات لعملة ؟ (ب) الحصول على 3 رقم ستة فى 4 رميات لنردة
الإجابة (أ) 0.23 (ب) 0.0154321 .

٣ – ٥٩ إذا كان 20% من الطلاب الذين يلتحقون بالجامعة يتركونها قبل إتمام الدراسة بها ، أوجد احتمال أنه من بين 20 طالباً تم اختيارهم عشوائياً من بين العدد الكبير جداً من الملتحقين بالجامعة ، أن 3 منهم سوف يتركون الجامعة قبل إتمام الدراسة .
(ب) إذا كان 90% من المصابيح الكهربائية المنتجة فى أحد المصانع مقبولة ، ما احتمال أن يكون من بين 10 مصابيح مختارة عشوائياً من إنتاج كبير جداً للمصنع ، 8 منها مقبولة .
الإجابة (أ) 0.206 (ب) 0.1937

٣ – ٦٠ احسب القيمة المتوقعة والانحراف المعيارى وحدد تماثل أو عدم تماثل التوزيع الاحتمالى
(أ) للمسألة (٣ – ٥٨ (أ)) ، (ب) للمسألة (٣ – ٥٩ (أ)) ، (جـ) للمسألة (٣ – ٥٩ (ب))
الإجابة (أ) $E(X) = 3$ صورة ، $SDX = 1.22$ صورة ، والتوزيع متماثل (ب) $E(X) = 4$ طالباً ، $SD\ X = 1.79$ طالباً ، والتوزيع ملتو لليمين (جـ) $E(X) = 9$ مصابيح $SD\ X = 0.95$ مصباح ، والتوزيع ملتو لليسار .

٣ – ٦١ ما احتمال اختيار (أ) سيدتين فى عينة من 5 أفراد مسحوبة عشوائياً بدون إحلال من مجموعة من 9 أفراد منهم 4 سيدات
(ب) ثمانى رجال فى عينة من 10 أفراد مسحوبة عشوائياً بدون إحلال من مجتمع مكون من 1,000 فرد نصفهم رجال ؟
الإجابة (أ) حوالى 0.71 (باستخدام التوزيع فوق الهندسى) (ب) حوالى 0.0439 (باستخدام تقريب ذى الحدين لاحتمالات التوزيع فوق الهندسى).

توزيع بواسون :

٣ – ٦٢ تشير الخبرة الماضية إلى أن هناك 2 حادثة مرور عند تقاطع ما أسبوعياً . ما احتمال (أ) 4 حوادث خلال أسبوع مختار عشوائياً ؟
(ب) عدم وقوع حوادث ؟ (جـ) ما القيمة المتوقعة والانحراف المعيارى للتوزيع ؟
الإجابة (أ) حوالى 0.36 (ب) حوالى 0.14 (جـ) $E(X) = \lambda = 2$ حادثة ، $SDX = \sqrt{\lambda} = 1.41$ حادثة .

٣ — ٦٣ تظهر الخبرة الماضية أن 0.003 من القوة العاملة القومية يصابون بمرض خطير خلال عام . فإذا اختير 1,000 شخص عشوائياً من بين القوة العاملة القومية :

(أ) ماهى القيمة المتوقعة لعدد العاملين الذين سوف يمرضون خلال العام ؟

(ب) ما احتمال أن 5 عاملين سوف يمرضون خلال العام ؟

الإجابة (أ) 3 عاملين (ب) حوالى 0.1 (باستخدام بواسون كتقريب لذى الحدين)

التوزيعات الاحتمالية المتصلة : التوزيع الطبيعى :

٣ — ٦٤ اعط الصيغ التالية :

(أ) احتمال أن يقع المتغير المتصل X بين X_1، X_2 .

(ب) التوزيع الطبيعى (جـ) القيمة المتوقعة والتباين للتوزيع الطبيعى . (د) التوزيع الطبيعى القياسى .

(هـ) ماهو متوسط وتباين التوزيع الطبيعى القياسى ؟

الإجابة (أ) $P(X_1 < X < X_2) = \int_{X_1}^{X_2} f(x)\,dX$ (ب) $f(X) = (1/\sqrt{2\pi\sigma^2})\exp\{-(1/2)[(X-\mu)/\sigma]^2$

(جـ) $= \int_{-\infty}^{\infty} Xf(X)\,dX$ and $\sigma^2 = \int_{-\infty}^{\infty}[X - E(X)]^2 f(X)\,dX$ (د) $f(X) = (1/\sqrt{2\pi})\exp[-(1/2)z^2]$

(هـ) $E(X) = \mu = 0$ and $\sigma^2 = 1$

٣ — ٦٥ أوجد المساحة تحت منحنى التوزيع الطبيعى القياسى (أ) بين $z = \pm 1.64$ (ب) بين $z = \pm 1.96$ (جـ) بين $z = \pm 2.58$ (د) بين $z = 0.90$ و $z = 2.10$ (هـ) إلى اليسار من $z = 0.90$ (و) إلى اليمين من $z = 2.10$ (ز) إلى اليسار من $z = 0.90$ وإلى اليمين من z 2.10

الإجابة (أ) 0.899 أو 89.90% (ب) 0.95 (جـ) 0.9902 (د) 0.1662 (هـ) 0.8159 (و) 0.0179 (ز) 0.8338

٣ — ٦٦ متغير عشوائى يتبع التوزيع الطبيعى مع $\mu = 67$ و $\sigma = 3$. ما احتمال أن هذا المتغير العشوائى سيأخذ القيمة (أ) بين 67 و70 ؟ (ب) بين 60 و 70 ؟ (جـ) بين 60 و 65 ؟ (د) أقل من 60 ؟ (هـ) أكبر من 65 ؟

الإجابة (أ) 0.3413 أو 34.13% (ب) 0.8334 (جـ) 0.2415 (د) 0.0099 (هـ) 0.7486

٣ — ٦٧ الوسط الحسابى لأوزان مجموعة كبيرة من الناس هو 180 رطلا والانحراف المعيارى 15 رطلا . إذا كانت الأوزان تتبع التوزيع الطبيعى ، أوجد احتمال أن شخصاً تم اختياره عشوائياً من المجموعة سوف يزن (أ) بين 160 و180 رطلا (ب) أعلى من 200 رطلا (جـ) أقل من 150 رطلا .

الإجابة (أ) 0.4082 أو 40.82% (ب) 0.0918 (جـ) 0.0228

٣ — ٦٨ تتبع درجات اختبارات الذكاء لمتطوعى الجيش فى سنة ما التوزيع الطبيعى مع $\mu = 110$ و $\sigma = 10$. ويريد الجيش أن يعطى تدريباً متقدماً لأعلى 25 % فى درجات اختبارات الذكاء . ماهى أقل درجة فى اختبارات الذكاء التى تقبل لحضور التدريب المتقدم ؟

الإجابة 117 لأقرب رقم صحيح

٣ — ٦٩ تشير الخبرة الماضية إلى أن 60% من الطلاب الملتحقين بالكليات يحصلون على مؤهلاتهم . باستخدام (أ) توزيع ذى الحدين ، (ب) التوزيع الطبيعى كتقريب لذى الحدين ، أوجد احتمال أنه من بين 30 طالباً مختارين عشوائياً من الملتحقين حديثاً سوف يحصل أكثر من 20 طالباً على المؤهل .

الإجابة (أ) 0.1762 (ب) 0.1762

٣ — ٧٠ فى المتوسط ، تمر 10 سيارات فى الدقيقة من أمام كشك تحصيل رسوم المرور خلال ساعة الذروة . باستخدام (أ) توزيع بواسون ، (ب) التوزيع الطبيعى كتقريب لبواسون ، أوجد احتمال أن أقل من 6 سيارات سوف تمر أمام الكشك خلال دقيقة مختارة عشوائياً .

الإجابة (أ) 0.0671 أو % 6.71 (ب) 0.0778 أو 7.74%

٣ — ٧١ تنتج عملية صناعية فى المتوسط 2 وحدة معيبة فى الساعة . ما احتمال أنه بعد الحصول على وحدة معيبة : (أ) أن تمر ساعة قبل الحصول على الوحدة المعيبة التالية ؟ (ب) أن تمر نصف ساعة ؟ (جـ) أن تمر خمسة عشرة دقيقة ؟ (د) ما القيمة المتوقعة والانحراف المعيارى لهذا التوزيع ؟

الإجابة (أ) 0.13534 أو 13.53% (ب) 0.36788 (جـ) 0.60653 (د) $E(T) = \sigma = 1/\lambda = 1/2\,\mathrm{h}$ لكل وحدة معيبة

٣ — ٧٢ متوسط تقديرات طالبة أعلى من المتوسط الحسابى فى مدرستها بعدد 3 انحراف معيارى . ماعدد طلاب المدرسة الذين لهـ . (أ) متوسط درجات أعلى ؟ (ب) متوسط درجات أقل ؟

الإجابة (أ) 0.11 أو 11% (باستخدام نظرية أو متباينة تشبتشيف) (ب) 0.89 أو 89%

الفصل الرابع

الاستدلال الاحصائى : التقدير

٤ــ١ المعاينة

الاستدلال الاحصائى واحد من أكثر جوانب عملية اتخاذ القرارات أهمية وحيوية فى الاقتصاد والأعمال والعلوم . ويتعلق الاستدلال الإحصائى بالتقدير واختبار الفروض (الفصل الخامس) . والتقدير هو عملية استنتاج أو تقدير أحد معالم المجتمع (مثل الوسط الحسابى أو الانحراف المعيارى) من الإحصاء المناظر والخاص بعينة مسحوبة من المجتمع .

ولكى يكون التقدير (واختبار الفروض) سليما ، ينبغى أن يبنى على عينة ممثلة للمجتمع . ويمكن تحقيق ذلك بالمعاينة العشوائية حيث يكون لكل مفردة فى المجتمع فرصة متكافئة للدخول فى العينة .

مثال ١ — يمكن الحصول على عينة عشوائية من 5 من بين 80 عاملا فى مصنع بتسجيل إسم كل منهم على قصاصة من الورق ، وخلط القصاصات جيداً ، والتقاط خمس منها عشوائياً . وكطريقة أقل تعقيداً يمكن استخدام جدول الأرقام العشوائية (ملحق ٤) . ولاستخدام هذه الطريقة ، نعين أولا رقماً من 1 إلى 80 لكل عامل . ثم نبدأ عند نقطة عشوائية (وليكن ، من العمود الثالث والصف الحادى عشر) فى ملحق ٤ ، ونقرأ 5 أرقام (كل رقمين معاً) أما أفقياً وإما رأسياً (مع حذف كل الأعداد التى تزيد عن 80) . على سبيل المثال ، بالقراءة رأسياً نحصل على 13 ، 54 ، 19 ، 59 ، 71 .

٤ــ٢ توزيع المعاينة للمتوسط

إذا أخذنا عينات متكررة من مجتمع ما وقنا بقياس متوسط لكل عينة ، فإننا نجد أن معظم هذه المتوسطات $\bar{X}_s$ ، تختلف عن بعضها البعض ، ويسمى التوزيع الاحتمالى لمتوسطات العينات هذه « توزيع المعاينة للوسط » . ولكن توزيع المعاينة للوسط له أيضاً وسط ، يعبر عنه بالرمز $\mu_{\bar{X}}$ ، وانحراف معيارى أو خطأ معيارى $\sigma_{\bar{X}}$.

وهناك نظريتان هامتان تربطان بين توزيع المعاينة للوسط والمجتمع الأصلى .

نظرية ١ — إذا أخذنا عينات متكررة حجمها n من مجتمع ما :

$$\mu_{\bar{X}} = \mu \qquad (٤ - ١)$$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \quad \text{or} \quad \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} \qquad \text{و} (٤ - ٢ أ ، ب)$$

حيث تستخدم معادلة (٤ — ٢ ب) للمجتمعات المحدودة ذات الحجم N عندما تكون $n \geq 0.05\,N$ (أنظر المسألة ٤ — ٥ (ب) .

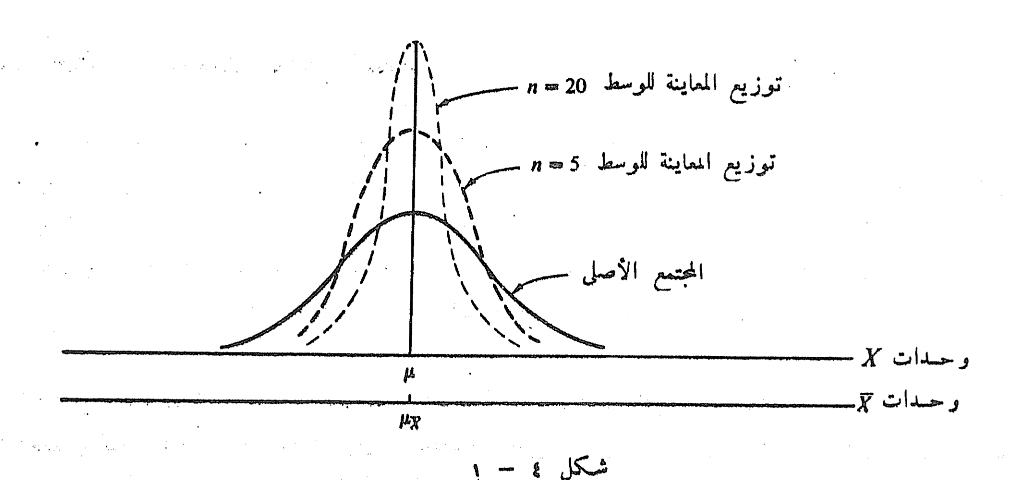
نظرية ٢ — مع تزايد حجم العينات (أى عندما $n \to \infty$) فإن توزيع المعاينة للمتوسط يقترب من التوزيع الطبيعى بصرف النظر عن شكل المجتمع الأصلى . ويعتبر التقريب جيداً عندما تكون $n \geq 30$. هذه هى نظرية النهاية المركزية .

ويمكن إيجاد احتمال أن يكون الوسط $\bar{X}$ لعينة عشوائية داخل فترة معينة ، بحساب قيم z للفترة ، حيث

$$z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} \qquad (٤ - ٣)$$

ثم الكشف عن هذه القيم في ملحق ٣ ، كما سبق شرحه في قسم ٣ — ٥ .

مثال ٢ — في شكل ٤ — ١ ، متوسط توزيع المعاينة للمتوسط $\mu_{\bar{X}}$ ، يساوي متوسط المجتمع μ بصرف النظر عن حجم العينة n ، ولكن ، كلما كبرت n ، كلما صغر الخطأ المعياري للمتوسط $\sigma_{\bar{X}}$ فإذا كان توزيع المجتمع هو التوزيع الطبيعي ، فإن توزيع المعاينة للوسط هو أيضاً التوزيع الطبيعي ، حتى للعينات الصغيرة . وطبقاً لنظرية النهاية المركزية ، حتى إذا كان المجتمع غير طبيعي التوزيع ، فإن توزيع المعاينة للمتوسط يكون طبيعياً تقريباً عندما تكون $n \geq 30$.



شكل ٤ — ١

مثال ٣ — افترض أن المجتمع يتكون من 900 عنصر بوسط حسابي 20 وحدة وانحراف معياري 12 وحدة . الوسط الحسابي والانحراف المعياري لتوزيع المعاينة لوسط عينة حجمها 36 هما

$$\mu_{\bar{X}} = \mu = 20 \text{ وحدة}$$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{12}{\sqrt{36}} = 2$$

لو كانت n تساوي 64 بدلا من 36 (بحيث $n > 0.05\,N$) ، فإن

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{12}{\sqrt{64}}\sqrt{\frac{900-64}{900-1}} = \frac{12}{8}\sqrt{\frac{836}{899}} = (1.5)(0.96) = 1.44$$

بدلا من $\sigma\bar{X} = 1.5$ بدون معامل التصحيح للمجتمعات المحدودة .

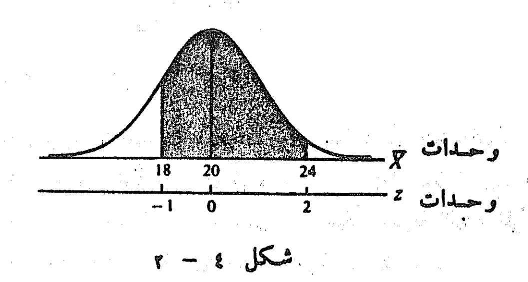
مثال ٤ — يمكن حساب احتمال أن يقع وسط عينة عشوائية $\bar{X}$ حجمها 36 مأخوذة من مجتمع مثال ٣ بين 18 و 24 كما يلي :

$$z_1 = \frac{\bar{X}_1 - \mu_{\bar{X}}}{\sigma_{\bar{X}}} = \frac{18-20}{2} = -1 \qquad \text{و} \qquad z_2 = \frac{\bar{X}_2 - \mu_{\bar{X}}}{\sigma_{\bar{X}}} = \frac{24-20}{2} = 2$$

بالبحث مقابل z_1 و z_2 في ملحق ٣ ، نحصل على :

$$P(18 < \bar{X} < 24) = 0.3413 + 0.4772 = 0.8185, \text{ or } 81.85\%$$

أنظر شكل ٤ - ٢



شكل ٤ - ٢

٤-٣ التقدير باستخدام التوزيع الطبيعى

يمكننا الحصول على تقدير لأحد معالم المجتمع إما بنقطة وإما بفترة . فالتقدير بنقطة عبارة عن عدد واحد . ويكون هذا التقدير بنقطة غير متحيز إذا كانت القيمة المتوقعة أو القيمة الوسطى للاحصاء المناظر ، عند تكرار المعاينة المشوائية ، مساوية لمعلمة المجتمع . فمثلا $\bar{X}$ هى تقدير (بنقطة) غير متحيز للمعلمة μ لأن $\mu_{\bar{X}} = \mu$ حيث $\mu_{\bar{X}}$ هى القيمة المتوقعة للمتوسط $\bar{X}$. أما الانحراف المعيارى s للعينة (كما هو معرف فى المعادلتين (٢ - ١٠ ب) ، (٢ - ١١ ب)) فهو تقدير غير متحيز للمعلمة σ (أنظر المسألة ٤ - ١٣ ب) ، والنسبة فى العينة p هى تقدير غير متحيز للمعلمة p (وهى نسبة المفردات التى لها خاصية معينة فى المجتمع كله) .

أما التقدير بفترة فيشير إلى مدى من القيم مقرونا باحتمال أن يضم هذا المدى (الفترة) معلمة المجتمع غير المعروفة ، ويسمى هذا الاحتمال مستوى الثقة . وبمعلومية الانحراف المعيارى للمجتمع أو تقديره ، وإذا علم أن توزيع المجتمع طبيعى أو علم أن العينة العشوائية تساوى أو تزيد عن 30 يمكننا إيجاد فترة الثقة 95% لوسط المجتمع غير المعروف كالآتى :

$$P(\bar{X} - 1.96\sigma_{\bar{X}} < \mu < \bar{X} + 1.96\sigma_{\bar{X}}) = 0.95 \qquad (٤ - ٤)$$

وتنص هذه المعادلة على أنه فى المعاينة العشوائية المتكررة ، نتوقع أن 95 فترة من 100 فترة كالتى فى معادلة ٤ -٤ تحتوى على معلمة المجتمع غير المعلومة وأن فترة الثقة التى لدينا (المبنية على عينة واحدة) هى واحدة من هؤلاء .

ويمكن تكوين فترة ثقة لنسبة المجتمع بأسلوب مماثل (أنظر مثال ٧) حيث

$$\mu_p = \frac{\mu}{n} = p \quad \text{(نسبة النجاح فى المجتمع)} \qquad (٤ - ٥)$$

$$\sigma_p = \sqrt{\frac{p(1-p)}{n}} \quad \text{(الخطأ المعيارى للنسبة)} \qquad (٤ - ٦)$$

مثال ٥ - أخذت عينة عشوائية حجمها 144 بوسط مقداره 100 وانحراف معيارى مقداره 60 من مجتمع حجمه 1000 . فترة الـ 95 % ثقة لوسط المجتمع غير المعلوم هى

$$\mu = \bar{X} \pm 1.96\sigma_{\bar{X}} \quad \text{حيث} \quad n > 30$$

$$= \bar{X} \pm 1.96\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} \quad \text{حيث} \quad n > 0.05N$$

$$= 100 \pm 1.96\frac{60}{\sqrt{144}}\sqrt{\frac{1,000-100}{1,000-1}} \quad \text{باستخدام } s \text{ كتقدير للمعلمة } \sigma$$

$$= 100 \pm 1.96(5)(0.95)$$

$$= 100 \pm 9.31$$

أى أن μ تقع بين 90.89 و 109.11 بدرجة ثقة 95% . وكثيرا ماتستخدم أيضاً درجات الثقة 90 و 99% وهى مناظرة لقيمة $z = 1.64$ و $z = 2.58$ على الترتيب (أنظر ملحق ٣) .

مثال ٦ – يرغب مدير فى تقدير متوسط عدد الدقائق التى يأخذها العمال لإنجاز عملية صناعية معينة فى حدود 3 ± دقيقة وبدرجة ثقة 90% . ويعلم المدير من خبرته الماضية أن الانحراف المعيارى σ هو 15 دقيقة . الحد الأدنى لحجم العينة المطلوب (n>30) يمكن إيجاده كالآتى :

$$z = \frac{\bar{X}-\mu}{\sigma_{\bar{X}}}$$

$$z\sigma_{\bar{X}} = \bar{X}-\mu$$

$$1.64\frac{\sigma}{\sqrt{n}} = \bar{X}-\mu \quad \text{بافتراض } n < 0.05N$$

$$1.64\frac{15}{\sqrt{n}} = 3 \quad \text{حيث أن خطأ التقدير } \bar{X}-\mu, \text{ is 3 min}$$

$$1.64\frac{15}{3} = \sqrt{n}$$

$$n = 67.24, \text{ أو } 68 \text{ مقرب للرقم الصحيح الأعلى}$$

مثال ٧ – وجدت إدارة التعليم لإحدى الولايات ، أن فى عينة من 100 شخص مختارين عشوائياً من بين الملتحقين بالجامعات 40% منهم قد حصلوا على درجات جامعية . لإيجاد فتر الـ 99% ثقة لنسبة الحاصلين على درجات جامعية من بين جميع الملتحقين بالجامعة ، نمضى كما يلى . نلاحظ أولا أن المشكلة تتعلق بالتوزيع ذى الحدين (أنظر قسم ٣–٣) . وحيث أن 30 < n ، وكلا من $np > 5$ و $n(1-p) > 5$ فإن التوزيع ذا الحدين يقترب من التوزيع الطبيعى الأبسط فى الاستخدام (أنظر قسم ٣ – ٥) . فيكون

$$p = \bar{p} \pm z\sigma_p$$

$$p = \bar{p} \pm z\sqrt{\frac{p(1-p)}{n}} \quad \text{بافتراض } n < 0.05N$$

$$= 0.4 \pm 2.58\sqrt{\frac{(0.4)(0.6)}{100}} \quad \text{باستخدام } \bar{p} \text{ كتقدير للنسبة } p$$

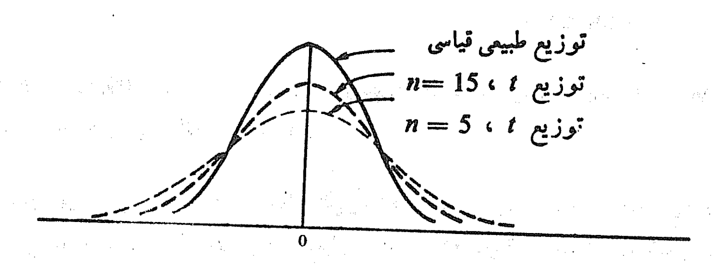
$$\cong 0.4 \pm 2.58(0.05)$$

$$\cong 0.4 \pm 0.13$$

إذن p تقع بين 0.27 و 0.53 بمستوى ثقة 99% .

٤-٤ فترات الثقة للمتوسط باستخدام توزيع t.

عندما يكون التوزيع طبيعياً ولكن σ غير معلومة و $n < 30$ ، فإننا لا نستطيع استخدام التوزيع الطبيعى لتحديد فترات الثقة لمتوسط المجتمع غير المعلوم ، ولكن يمكننا استخدام توزيع t . هذا التوزيع متماثل حول متوسط الصفر ولكنه منبسط عن التوزيع الطبيعى القياسى ، ولهذا فان جزءاً أكبر من مساحته تقع عند الأطراف . وبينما يوجد توزيع طبيعى قياسى واحد ، فإن هناك توزيعاً t مختلفاً لكل حجم للعينة n . ولكن مع تزايد n ، فإن توزيع t يقترب من التوزيع الطبيعى القياسى (أنظر شكل ٤ - ٣) إلى أن تكون $n \geqq 30$ ، وعندئذ يتساويان تقريباً .



شكل ٤ - ٣

ويعطى ملحق ٥ قيم t التى على يمينها تمثل المساحة تحت المنحنى 10% ، 5% ، 2.5% ، 1% و 0.5% من المساحة الكلية تحت المنحنى لدرجات الحرية المختلفة . وتعرف درجات الحرية df فى هذه الحالة بأنها $n-1$ (أو حجم العينة ناقص 1 ، وإذا كنا نرغب فى تقدير معلمة واحدة μ) . وفترة الثقة 95% لوسط المجتمع غير المعلوم عند استخدام توزيع t هى :

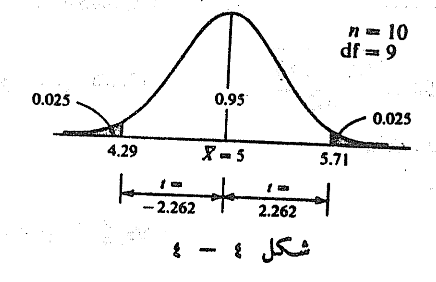
$$P\left(\bar{X} - t\frac{s}{\sqrt{n}} < \mu < \bar{X} + t\frac{s}{\sqrt{n}}\right) = 0.95 \qquad (٤ - ٧)$$

حيث تشير t إلى قيمة t التى تقع عندها 2.5% من المساحة الكلية للمنحنى عند كل طرف (عند درجات الحرية المستخدمة) ، وتستخدم $s/\sqrt{n}$ بدلا من $\sigma_{\bar{X}} = \sigma / \sqrt{n}$.

مثال ٨ - سحبت عينة عشوائية من $n=10$ بطارية فلاش بمتوسط $\bar{X}=5$ ساعة ، وانحراف معيارى للعينة $s = 1$ ساعة من خط إنتاج من المعروف أنه ينتج بطاريات عمرها موزع طبقاً للتوزيع الطبيعى . لإيجاد فترة الـ 95% ثقة للمتوسط غير المعلوم لعمر البطاريات فى المجتمع كله ، فإننا نوجد أولا قيمة $\pm t_{0.025}$ والتى تكون معها 2.5% من المساحة عند الأطراف لدرجات حرية $n-1=9$ ونحصل على هذه القيمة من ملحق ٥ بالتحرك تحت عمود 0.025 حتى درجات حرية 9 والقيمة التى نحصل عليها هى 2.262 إذن

$$\mu = \bar{X} \pm 2.262\frac{s}{\sqrt{n}} = 5 \pm 2.262\frac{1}{\sqrt{10}} \cong 5 \pm 2.262(0.316) \cong 5 \pm 0.71$$

وتقع μ بين 4.29 و 5.71 ساعة بدرجة ثقة 95% (أنظر شكل ٤-٤) . وعندما تكون $n < 30$ والتوزيع غير طبيعى ، فيجب استخدام نظرية تشبتشيف (أنظر المسألة ٤ - ٢٧) .



شكل ٤ - ٤

مسائل محلولة

٤ — ١ (أ) ما المقصود بالاستدلال الإحصائى ؟ وما وظيفته وما أهميته ؟ (ب) ما العلاقة بين المعلمة والإحصائية وماذا يقصد بهما ؟

(جـ) ما المقصود بالتقدير ؟ باختبار الفروض ؟

(أ) الاستدلال الإحصائى هو عملية استنتاج عن المجتمع من المعلومات التى تقدمها العينات . ويشمل المجتمع مجموعة العناصر (أشخاص، أجزاء تنتجها ماكينة ، سيارات تمر أمام نقاط مراقبة ، الخ) التى نقوم بوصفها. العينة هى الجزء المختار من المجتمع . وتحليل المجتمع كله قد يكون مستحيلا (إذا كان المجتمع غير محدود) وقد يدمر كل الإنتاج (كما فى حالة اختبار المصابيح الكهربائية المنتجة) ، وقد يكون بالغ التكلفة . ويمكن التغلب على هذه المشاكل بأخذ عينة (ممثلة) من المجتمع ، وعمل استدلالات عن المجتمع من العينة .

(ب) المعلمة هى خاصية وصفية (مثل الوسط والانحراف المعيارى) لمجتمع ما . أما الإحصائية فهى خاصية وصفية لعينة . وفى الاستدلال الإحصائى فإننا نقوم بعمل استدلالات عن المعالم من الإحصائيات المناظرة لها .

(جـ) الاستدلال الإحصائى نوعان : التقدير واختبار الفروض . والتقدير هو عملية استنتاج أو تقدير المعلمة من الإحصائية المناظرة . فمثلا ، يمكننا تقدير الوسط والانحراف المعيارى لمجتمع ما من الوسط والانحراف المعيارى لعينة مسحوبة من المجتمع . اختبار الفروض هو تحديد ما إذا كنا نقبل أو نرفض فرضاً ما عن معلمة على ضوء معلومات العينة . وتتناول التقدير فى هذا الفصل ، واختبارات الفروض فى الفصل الخامس .

٤ — ٢ ما المقصود بالمعاينة العشوائية ؟ ما أهميتها ؟

المعاينة العشوائية هى أسلوب معاينة يجعل لكل مفردة فى المجتمع فرصة متكافئة فى الدخول فى العينة . وتضمن المعاينة العشوائية أن تكون العينة ممثلة . وهناك عدة أنواع للمعاينة العشوائية . فى المعاينة العشوائية البسيطة ، لا يكون فقط لكل مفردة ولكن أيضاً لكل عينة نفس الفرصة أن تختار . فى المعاينة المنتظمة تسحب المفردات من المجتمع على فترات متساوية ، زمنا ، أو ترتيبياً أو مكانياً . (مثل اختيار اسم بعد كل مائة اسم فى دليل التليفون) . ويمكن أن تتحيز المعاينة المنتظمة بسهولة ، إذا قيست كمية فضلات المنازل كل يوم اثنين (حيث تشمل فضلات نهاية الأسبوع) . وفى المعاينة الطبقية والعنقودية ، يقسم المجتمع إلى طبقات (مثلا طبقاً للسن) وإلى مجموعات (مثل البلوكات السكنية فى المدينة) ثم تسحب نسب متساوية من كل طبقة أو مجموعة . وتستخدم المعاينة الطبقية عندما يكون الاختلاف داخل كل طبقة صغيرا بالنسبة للاختلاف بين المجموعات . وتستخدم المعاينة العنقودية فى الحالة العكسية . وسوف نفترض ، فيما يلى ، استخدام المعاينة العشوائية البسيطة . ويمكن أن تكون المعاينة من مجتمع محدود (أو سحب ورقة من مجموعة أوراق لعب بدون إحلال) أو من مجتمع غير محدود (كالتقاط أجزاء منتجة من خلال عملية مستمرة أو سحب أوراق من مجموعة أوراق لعب مع الإحلال) .

٤ — ٣ (أ) كيف يمكن الحصول على عينة عشوائية ؟ (ب) باستخدام جدول أعداد عشوائية ، اشتق عينة عشوائية من عشرة أفراد تغيبوا عن العمل بسبب المرض فى أحد الأيام من بين 95 موظفاً فى مصنع . (جـ) اشتق عينة عشوائية من 12 مفردة من بين 240 جزءاً أنتجتها ماكينة خلال الساعة الأولى من تشغيلها .

(أ) يمكن الحصول على عينة عشوائية (١) باستخدام كمبيوتر مبرمج لتجميع الأرقام ، (٢) من جدول أعداد عشوائية ، (٣) بتعيين رقم لكل عنصر فى المجتمع ، وتسجيل كل رقم على قصاصة منفصلة ، وخلط القصاصات جيداً ، ثم سحب عدد من القصاصات يعادل عدد العناصر المطلوبة من العينة . والطريقة الأخيرة لحصول على عينة عشوائية معقدة جداً فى حالة المجتمعات الكبيرة وقد لا تعطى عينة ممثلة بسبب صعوبة خلط القصاصات جيداً .

(ب) لحصول على عينة عشوائية من 10 مفردات من بين 95 موظفاً ، فإننا نعين لكل موظف رقما من 1 إلى 95 ثم نرجع ملحق ٤ (جدول الأعداد العشوائية) . ويحتوى ملحق ٤ على 1600 رقماً فى مجموعات من 5 أرقام ثم توليدها بعملية عشوائية تماماً بحيث أن كل رقم وكذلك تتابع الأرقام يكون لها احتمالات متساوية فى الحدوث . بالبدء عند نقطة اختيارية فى ملحق ٤ (مثلا العمود الرابع عشر والصف الخامس) ، وقراءة 10 أرقام من حدين (وليكن ، رأسياً مع حذف كل الأرقام أكبر من 95) نحصل على العينة العشوائية الآتية :

60 ، 39 ، 4 ، 34 ، 76 ، 43 ، 52 ، 15 ، 8 ، و 95

(جـ) لنبدأ مثلا ، من العمود الثالث والسطر الثامن فى ملحق ٤ ، ونتحرك أفقياً وتقرأ 8 أعداد (كل منها مكون من ثلاثة حدود مع حذف الأعداد التى تزيد عن 240 ، فنحصل على العينة العشوائية التالية :

215 ، 182 ، 51 ، 9 ، 127 ، 177 ، 53 ، و186

(وقد حصلنا على الأعداد الأربعة الأخيرة من السطر التاسع بعدما وصلنا إلى نهاية السطر الثامن) .

توزيع المعاينة للوسط :

٤ – ٤ (أ) ماذا يقصد بتوزيع المعاينة للوسط ، وكيف يتم الحصول عليه ؟ (ب) ماذا يقصد بالمتوسط والخطأ المعيارى لتوزيع المعاينة للوسط ؟

(أ) إذا أخذنا عينات عشوائية متكررة (أو كل العينات الممكنة) ، كل منها من حجم n ، من مجتمع من القيم للمتغير X وأوجدنا متوسط كل عينة $\bar{X}$ ، فإننا نجد أن معظم متوسطات العينات تختلف عن بعضها البعض . ويسمى التوزيع الاحتمالى لمتوسطات العينات هذه **توزيع المعاينة النظرى للوسط** . وبالمثل ، يمكن الحصول على توزيع المعاينة النظرى للنسبة ، وللفرق بين متوسطين ، وللفرق بين نسبتين ، فمثلا كان من الممكن إيجاد نسبة الوحدات المعيبة فى كل عينة ، والحصول على توزيع المعاينة النظرى لنسبة الوحدات المعيبة . وللتبسيط . فسوف نتناول فى هذا القسم توزيع المعاينة للوسط فقط .

(ب) كما فى التوزيعات الاحتمالية الأخرى (أنظر أقسام ٣ – ٣ إلى ٣ – ٥) ، يمكن توصيف توزيع المعاينة النظرى باستخدام الوسط والانحراف المعيارى . ويعطى متوسط توزيع المعاينة بالرمز $\mu_{\bar{X}}$ (وتقرأ « ميو برمز الدليل $\bar{X}$). وهذه هى متوسط رقم $\bar{X}$ والتى يجب تمييزها عن μ (متوسط المجتمع) . ويعطى الانحراف المعيارى لتوزيع المعاينة للوسط بالرمز $\sigma_{\bar{X}}$ (وتقرأ سيجما برمز الدليل $\bar{X}$) . وهذا هو الانحراف المعيارى لقيم $\bar{X}$ ويجب تمييزه عن σ (الانحراف المعيارى للمجتمع) . وكلما صغر $\sigma_{\bar{X}}$ ، كلما كان متوسط العينة $\bar{X}$ أكثر دقة كتقدير لمتوسط المجتمع غير المعلوم μ . ولهذا ، عادة ما يشار إلى $\sigma_{\bar{X}}$ **بالخطأ المعيارى** للوسط .

٤ – ٥ كيف يمكن إيجاد (أ) متوسط توزيع المعاينة للوسط $\mu_{\bar{X}}$ ؟ (ب) الانحراف المعيارى لتوزيع المعاينة للوسط أو الخطأ المعيارى $\mu_{\bar{X}}$ ؟

(أ) متوسط توزيع المعاينة النظرى للوسط $\mu_{\bar{X}}$ يساوى متوسط المجتمع الأصلى μ أى $\mu_{\bar{X}} = \mu$. لاحظ أنه لكى يكون هذا صحيحاً ، فإننا يجب إما أن نأخذ جميع العينات الممكنة من حجم n من المجتمع المحدود ، وإما ، إذا كنا نتعامل مع مجتمعات غير محدودة (أو مجتمعات محدودة مع الإحلال) فإنه يجب أن نستمر فى أخذ عينات متكررة من حجم n إلى مالا نهاية . بالإضافة ، فإن $\mu_{\bar{X}}$ تساوى يضاً $E(\bar{X})$ (أنظر المسألتين ٣ – ٢٠ ، ٣ – ٣١) .

(ب) الخطأ المعيارى للوسط $\sigma_{\bar{X}}$ هو الانحراف المعياري للمجتمع ، σ ، مقسوماً على الجذر التربيعى لحجم العينة ، $\sqrt{n}$ أى $\sigma_{\bar{X}} = \sigma/\sqrt{n}$. وبالنسبة للمجتمعات المحدودة من حجم N ، يجب استخدام معامل تصحيح وتكون $\sigma_{\bar{X}} = (\sigma/\sqrt{n})\sqrt{(N-n)/(N-1)}$. ولكن ، إذا كان حجم العينةصغيراً جداً بالنسبة لحجم المجتمع فإن $(N-n)/(N-1)$ تكون قريبة من 1 ويمكن إسقاطها من المعادلة . وقد جرت العادة على إهمال هذا الحد عندما تكون $n < 0.05\,N$. وبصرف النظر عن معامل التصحيح ، فإن $\sigma_{\bar{X}}$ ترتبط طردياً مع σ وعكسياً مع $\sqrt{n}$ (أنظر المعادلتين (٤ - ٢ أ ، ب)) . أى أن مضاعفة حجم العينة 4 مرات يرفع دقة $\bar{X}$ كتقدير للوسط μ بخفض $\sigma_{\bar{X}}$ إلى النصف . لاحظ أيضاً أن $\sigma_{\bar{X}}$ دائماً أصغر من σ . والسبب ، أن أوساط العينات ، كمتوسطات لمشاهدات العينة تظهر اختلافاً أو انتشاراً أقل من قيم المجتمع . بالإضافة كلما كبر حجم العينة ، كلما صغر حجم $\sigma_{\bar{X}}$ بالنسبة لحجم σ (أنظر شكل ٤ - ١) .

٤ - ٦ من مجتمع مكون من الأعداد الخمسة الآتية : 1 ، 3 ، 5 ، 7 و 9 . أوجد (أ) μ و σ (ب) توزيع المعاينة النظرى للوسط لحجم العينة 2 ، (جـ) $\mu_{\bar{X}}$ و $\sigma_{\bar{X}}$

(أ)

$$\mu = \frac{\sum X}{N} = \frac{1+3+5+7+9}{5} = \frac{25}{5} = 5$$

$$\sigma = \sqrt{\frac{\sum (X-\mu)^2}{N}} = \sqrt{\frac{(1-5)^2+(3-5)^2+(5-5)^2+(7-5)^2+(9-5)^2}{5}}$$

$$= \sqrt{\frac{16+4+0+4+16}{5}} = \sqrt{\frac{40}{5}} = \sqrt{8} \cong 2.83$$

(ب) توزيع المعاينة النظرى للوسط لحجم عينة 2 من مجتمع محدود ، عبارة عن متوسط كل العينات الممكنة التى يمكن الحصول عليها من المجتمع . عدد توافيق 5 أشياء مأخوذاً منها 2 فى كل مرة بدون نظر إلى الترتيب هى $5!/2!\,3! = 10$ (أنظر المسألة ٣-١٨) . وهذه العينات العشر هى: 1,3 ، 1,5 ، 1,7 ، 1,9 ، 3,5 ، 3,7 ، 3,9، 5,7 ، 5,9، 7,9 والمتوسطات $\bar{X}$ للعينات العشر السابقة هى 2 ، 3 ، 4 ، 5 ، 4 ، 5 ، 6 ، 7 ، 6 ، 8 ويعطى جدول ٤-١ التوزيع النظرى للمعاينة . لاحظ أن التغير أو انتشار متوسطات العينات (من 2 إلى 8) أقل من التغير أو انتشار القيم فى المجتمع الأصلى (من 1 إلى 9) ، تأكيداً للتقرير الوارد فى نهاية المسألة ٤ - ٥ (ب) .

جدول ٤ - ١ توزيع المعيانة النظرى للوسط

قيم المتوسط	النواتج الممكنة	احتمال الحدوث
2	2	0.1
3	3	0.1
4	4, 4	0.2
5	5, 5	0.2
6	6, 6	0.2
7	7	0.1
8	8	0.1
		Total 1.0

(جـ) بتطبيق نظرية ١ (قسم ٤ – ٢) ، $\mu_{\bar{X}} = \mu = 5$. وحيث أن حجم العينة 2 أكبر من 5% من حجم المجتمع (أى $n > 0.05N$) ،

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} = \frac{\sqrt{8}}{\sqrt{2}} \sqrt{\frac{5-2}{5-1}} = \sqrt{4}\sqrt{\frac{3}{4}} = \sqrt{3} \cong 1.73$$

٤ – ٧ بالنسبة لتوزيع المعاينة النظرى لمتوسط العينة المحسوب فى المسألة ٤ – ٦ (ب) ، (أ) أوجد الوسط والخطأ المعيارى للوسط باستخدام المعادلات الخاصة بوسط المجتمع ولانحرافه المعيارى كما فى أقسام ٢ – ٢ و ٢ – ٣ (ب) ماذا توضح الإجابات فى (أ)

(أ)
$$\mu_{\bar{X}} = \frac{\sum \bar{X}}{N} = \frac{2+3+4+5+4+5+6+6+7+8}{10} = \frac{50}{10} = 5$$

$$\sigma_{\bar{X}} = \sqrt{\frac{\sum(\bar{X}-\mu_{\bar{X}})^2}{N}}$$

$$= \sqrt{\frac{(2-5)^2+(3-5)^2+(4-5)^2+(5-5)^2+(4-5)^2+(5-5)^2+(6-5)^2+(6-5)^2+(7-8)^2+(8-5)^2}{10}}$$

$$= \sqrt{\frac{9+4+1+0+1+0+1+1+4+9}{10}} = \sqrt{\frac{30}{10}} = \sqrt{3} \cong 1.73$$

(ب) تؤكد إجابات (أ) النتائج السابق الحصول عليها فى تمرين ٤ – ٦ (جـ) بتطبيق نظرية ١ (قسم ٤ – ٢) ، أى $\mu_{\bar{X}} = \mu$: و $\sigma_{\bar{X}} = (\sigma/\sqrt{n})\sqrt{(N-n)/(N-1)}$ للمجتمع المحدود حيث $n > 0.05N$. لاحظ أننا أخذنا جميع العينات الممكنة من حجم 2 التى يمكن أخذها من مجتمع محدود به 5 مفردات . أما المعاينة من مجتمع غير محدود (أو من مجتمع محدود مع الاحلال) فإنه يتطلب أخذ عدد غير محدود من العينات العشوائية من حجم n من المجتمع الأصلى (واضح أنها مهمة غير ممكنة) . فإذا أخذنا فقط عدداً محدوداً من العينات العشوائية فإن نظرية ١ تنطبق فقط كتقريب (أى أن $\mu_{\bar{X}} \approx \mu$ و $\sigma_{\bar{X}} \approx \sigma/\sqrt{n}$) ، مع تحسن التقريب مع زيادة عدد العينات العشوائية المأخوذة . وفى هذه الحالة يشار إلى توزيع المعاينة لوسط العينات المتولد بتوزيع المعايينة التجريبى للوسط .

٤ – ٨ مجتمع مكون من 12,000 عنصر بوسط 100 وانحراف معيارى 60 . أوجد الوسط والخطأ المعيارى لتوزيع المعاينة للمتوسط عندما يكون حجم العينة (أ) 100 ، (ب) 900 .

(أ)
$$\mu_{\bar{X}} = \mu = 100$$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{60}{\sqrt{100}} = 6$$

(ب)
$$\mu_{\bar{X}} = \mu = 100$$

وحيث أن حجم العينة 900 أكبر من 5% حجم المجتمع ، فإن معامل التصحيح يجب أن يستخدم فى معادلة الخطأ المعيارى :

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} = \frac{60}{\sqrt{900}} \sqrt{\frac{12,000-900}{12,000-1}} = \frac{60}{30} \sqrt{\frac{11,100}{11,999}} \cong 2\sqrt{0.925} \cong 2(0.962) \cong 1.92$$

وبدون معامل التصحيح ، فإن $\sigma_{\bar{X}}$ تكون مساوية 2 بدلا من 1.92 .

٤ — ٩ (أ) ما شكل توزيع المعاينة النظرى للوسط إذا كان المجتمع الأصل طبيعياً ؟ إذا كان المجتمع الأصل غير طبيعى ؟
(ب) ما أهمية الإجابة على (أ) ؟

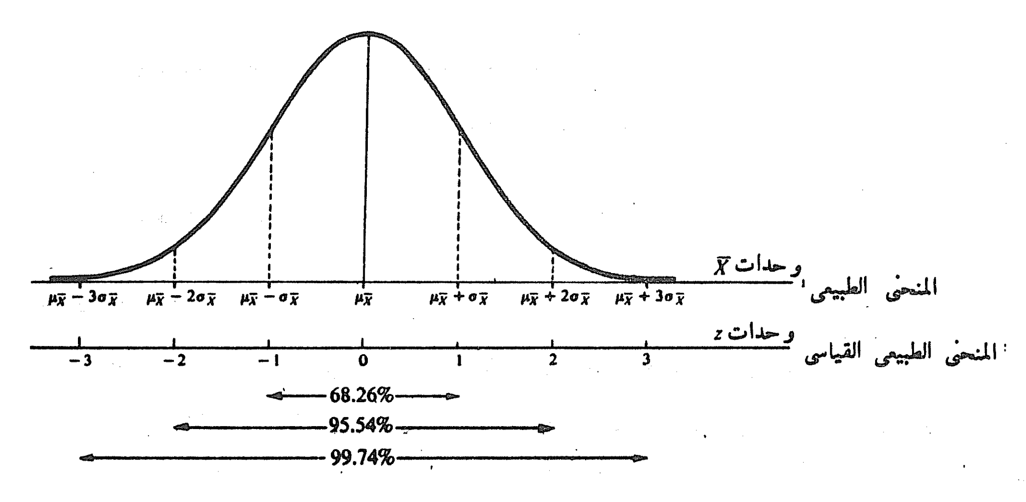
(أ) إذا كان توزيع المجتمع الأصل طبيعياً ، فإن توزيعات المعاينة النظرية للوسط تكون أيضاً طبيعية ، بصرف النظر عن حجم العينة . وطبقاً لنظرية النهاية المركزية ، حتى إذا لم يكن المجتمع الأصل طبيعياً ، فإن توزيعات المعاينة النظرية لمتوسط العينة يقترب من التوزيع الطبيعى مع تزايد حجم العينة (أى ، عندما $n \to \infty$) . ويكون التقريب جيداً للعينات عند حجم 30 على الأقل .

(ب) ربما تكون نظرية النهاية المركزية أهم نظرية فى الاستدلال الإحصائى . فهى تسمح لنا باستخدام إحصائيات العينة لعمل استنتاجات عن معلمات المجتمع بدون معرفة أى شئ عن شكل المجتمع الأصل . ونحن نقوم بذلك فى هذا الفصل وفى الفصل الخامس .

٤ — ١٠ كيف يمكننا حساب احتمال وقوع وسط عينة عشوائية داخل فترة معينة إذا كان توزيع المعاينة النظرى للوسط طبيعياً أو قريباً من الطبيعى ؟ كيف يختلف ذلك عن عملية إيجاد احتمال أن يأخذ متغير عشوائى توزيعه طبيعى قيمة داخل فترة معينة ؟
(ب) ارسم المنحنى الطبيعى باستخدام وحدات $\bar{X}$ ووحدات z وبين نسبة المساحة تحت المنحنى على بعد 1 ، 2 ، 3 انحراف معيارى من الوسط .

(أ) إذا كان توزيع المعاينة النظرى للوسط طبيعياً أو قريباً من الطبيعى ، فإنه يمكننا حساب احتمال أن يقع متوسط عينة عشوائية $\bar{X}$ داخل فترة معينة بحساب قيمة z المناظرة فى ملحق ٣ . وهذا يماثل ماسبق عمله فى قسم ٣—٥ عندما قدمنا المنحنى الطبيعى والمنحنى الطبيعى القياسى . الفرق الوحيد هو أننا نتعامل الآن مع توزيع المتوسطات $\bar{X}$s بدلا من توزيع المشاهدات Xs . وبالإضافة فإنه قبلا كانت $z = (X - \mu)/\sigma$ ، فأصبحت الآن $z = (\bar{X} - \mu_{\bar{X}})/\sigma_{\bar{X}} = (\bar{X} - \mu)/\sigma_{\bar{X}}$ حيث $\mu_{\bar{X}} = \mu$.

(ب) لدينا فى شكل ٤ — ٥ . منحنى طبيعى بوحدات $\bar{X}$ ومنحنى طبيعى قياسى بوحدات z . والمساحة تحت المنحنى على بعد 1 ، 2 و 3 انحراف معيارى هى 68.26 ، 95.54 و 99.7% على الترتيب . لاحظ التشابه الكبير وكذلك الاختلاف العام بين شكل ٤ — ٥ وشكل ٣ — ٤ .



شكل ٤ — ٥

٤ – ١١ أوجد احتمال أن يكون وسط عينة عشوائية من 25 عنصراً مأخوذة من مجتمع طبيعى بمتوسط 90 وانحراف معيارى 60 أكبر من 100 .

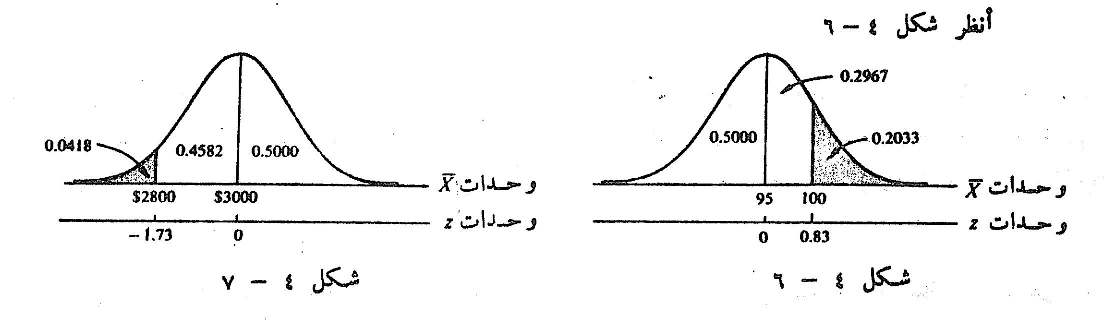
حيث أن المجتمع الأصلى موزع طبيعياً ، فإن توزيع المعاينة النظرى للوسط تكون أيضاً طبيعياً ويكون $\sigma_{\bar{X}} = \sigma/\sqrt{n}$ ، لأن $n < 0.05\,N$. عندما $\bar{X} = 100$ ،

$$z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{100 - 90}{60/\sqrt{25}} = \frac{10}{12} \cong 0.83$$

بالكشف عن هذه القيمة فى ملحق ٣ ، نحصل على

$$P(\bar{X} > 100) = 1 - (0.5000 + 0.2967) = 1 - 0.7967 = 0.2033, \text{ or } 20.33\%$$

أنظر شكل ٤ – ٦



شكل ٤ – ٦

شكل ٤ – ٧

٤ – ١٢ لدى بنك محلى صغير 1,450 حساب ادخار شخصى برصيد متوسط قدره 3,000 $ وانحراف معيارى 1,200 $. إذا أخذ البنك عينة عشوائية من 100 حساب ، ما احتمال أن متوسط المدخرات لهذه الحسابات المائة سيكون أقل من 2,800 $ ؟

حيث $n = 100$ ، فإن توزيع المعاينة النظرى للوسط يكون قريباً من الطبيعية ، ولكن حيث أن $n > 0.05N$ ، يجب استخدام معامل التصحيح لإيجاد $\sigma_{\bar{X}}$. عندما $\bar{X} = \$\ 2{,}800$

$$z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} = \frac{\bar{X} - \mu}{\dfrac{\sigma}{\sqrt{n}}\sqrt{\dfrac{N-n}{N-1}}} = \frac{2{,}800 - 3{,}000}{\dfrac{1{,}200}{\sqrt{100}}\sqrt{\dfrac{1{,}450 - 100}{1{,}450 - 1}}} = \frac{-200}{120\sqrt{\dfrac{1350}{1449}}} \cong \frac{-200}{120(0.965)} \cong -1.73$$

بالكشف مقابل $z = 1.73$ فى ملحق ٣ ، نحصل على :

$$P(\bar{X} < \$2{,}800) = 1 - (0.5000 + 0.4582) = 1 - 0.9582 = 0.0418, \text{ or } 4.18\%$$

أنظر شكل ٤ – ٧ .

التقدير باستخدام التوزيع الطبيعى :

٤ – ١٣ ماذا يعنى (أ) التقدير بنقطة ؟ (ب) مقدر غير متحيز ؟ (جـ) التقدير بفترة ؟

(أ) كثيراً ما يتم تقدير معامل المجتمع باستخدام إحصائيات العينة بسبب عوامل التكلفة والوقت والإمكانية . وتسمى إحصائية العينة المستخدمة فى تقدير معلمة المجتمع المقدر ، وتسمى القيمة المعينة المشاهدة تقدير . وعندما نعبر عن تقدير معلمة المجتمع بعدد واحد فإنه يسمى تقدير بنقطة فمثلا ، متوسط العينة $\bar{X}$ ، هو مقدر لوسط المجتمع μ ،

وقيمة مفردة للمتوسط $\bar{X}$ هو تقدير بنقطة لوسط المجتمع μ . وبالمثل فإنه يمكن استخدام الانحراف المعيارى للعينة s كمقدر للانحراف المعيارى للمجتمع σ ، والقيمة المفردة للانحراف المعيارى s كتقدير بنقطة للانحراف المعيارى للمجتمع σ . كذلك يمكن استخدام النسبة فى العينة $\bar{p}$ ، كمقدر للنسبة p ، فى المجتمع ، والقيمة المفردة للنسبة $\bar{p}$ يمكن استخدامها كتقدير للنسبة p (أى الجزء من المجتمع الذى له خواص معينة) .

(ب) يعتبر المقدر غير متحيز إذا أعطى توزيع المعاينة النظرى ، الناتج عن المعاينة العشوائية المتكررة من المجتمع ، إحصائية مساوية لمعلمة المجتمع . أو بعبارة أخرى ، فإن المقدر يكون غير متحيز إذا كانت قيمته المتوقعة (أنظر المسألتين ٣ — ٢٠ و ٣ — ٣١) مساوية لمعلمة المجتمع موضع التقدير . فمثلا $\bar{X}$ ، s (كما هى معرفة فى معادلات (٢ — ١٠ ب) و (٢ — ١١ ب) ، و $\bar{p}$ مقدرات غير متحيزة للمعالم μ ، σ و p على الترتيب . وهناك معايير هامة أخرى لما يعتبر مقدراً جيداً نناقشها فى قسم ٦ — ٤ .

(جـ) التقدير بفترة يشير إلى مدى القيم المستخدم لتقدير معلمة المجتمع غير المعلومة ، مع الاحتمال المناظر ، أو مستوى الثقة ، بأن تقع معلمة المجتمع غير المعلومة داخل هذه الفترة . وتعرف الفترة باسم فترة ثقة وهى تتمركز فى العادة حول تقدير بنقطة غير متحيز . فمثلا ، فترة الـ 95% ثقة للوسط μ .

$$P(\bar{X} - 1.96\sigma_{\bar{X}} < \mu < \bar{X} + 1.96\sigma_{\bar{X}}) = 0.95$$

ويسمى العددان المحددان لفترة الثقة باسم حدود الثقة . ولأن التقدير بفترة يعبر أيضاً عن درجة الدقة أو الثقة التى لدينا فى التقدير ، فإنه يمتاز عن التقدير بنقطة .

٤ — ١٤ عينة عشوائية من 64 مفردة وسطها 50 وانحرافها المعيارى 20 أخذت من مجتمع عدد مفرداته 800 (أ) أوجد تقدير بفترة لوسط المجتمع نكون معه واثقين 95% أن الفترة تتضمن وسط المجتمع . (ب) بماذا تخبرنا النتيجة فى (أ) ؟

(أ) حيث أن $n > 30$ ، فإننا نستخدم قيمة $z = 1.96$ من التوزيع القياسى الطبيعى لتكوين فترة الثقة 95% للمجتمع غير المعلوم ويمكننا استخدام s كتقدير للانحراف المعيارى σ غير المعلوم ، أى

(٤ — ٨) $\hat{\sigma} = s$ حيث تشير (^) إلى تقدير

(٤ — ٩ أ ، ب) $n > 0.05N$ عندما $\hat{\sigma}_{\bar{X}} = \frac{\hat{\sigma}}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$ أو $\hat{\sigma}_{\bar{X}} = \frac{\hat{\sigma}}{\sqrt{n}} = \frac{s}{\sqrt{n}}$

فى هذه المسألة

وتصبح (٢٢) $\hat{\sigma}_{\bar{X}} = \frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{20}{\sqrt{64}}\sqrt{\frac{800-64}{800-1}} \approx \frac{20}{8}0.96 \cong 2.4$

أى أن μ تقع بين حد الثقة الأدنى 45.3 وحد الثقة الأعلى 54.7 بدرجة ثقة 95% .

(ب) تخبرنا نتيجة (أ) أننا إذا أخذنا عينات عشوائية متكررة من المجتمع كلها من حجم $n = 64$ ، وأنشأنا فترات الثقة 95% لمتوسطات العينات ، فإن 95% من فترات الثقة هذه سوف تضم الوسط الحقيقى غير المعلوم للمجتمع . بافتراض أن فترة الثقة لدينا (المبنية على عينة عشوائية واحدة التى تم أخذها) هى واحدة من فترات الثقة 95% هذه التى تضم μ ، فإننا نأخذ المخاطرة المحسوبة بأننا على خطأ فى 5% من الحالات .

٤ - ١٥ عينة عشوائية من 25 مفرد لمتوسط 80 أخذت من مجتمع عدد مفرداته 1,000 توزيعه طبيعى بانحراف معيارى 30. أوجد فترات الثقة الآتية لوسط المجتمع غير المعلوم (أ) 90 % (ب) 95 % و (جـ) 99 % . (د) علام تدل الفروق فى نتائج (أ) ، (ب) و (جـ) ؟

(أ) حيث أن المجتمع يتبع التوزيع الطبيعى $\mu = \bar{X} \pm 1.64\sigma_{\bar{X}}$

(ب) حيث $n < 0.05\,N$ و σ معلومة . $\mu = \bar{X} \pm 1.64 \frac{\sigma}{\sqrt{n}}$

$$= 80 \pm 1.64 \frac{30}{\sqrt{25}}$$

$$= 80 \pm 1.64(6)$$

$$= 80 \pm 9.84$$

أى أن μ تقع بين 70.16 و 89.84 بمستوى ثقة 90% .

(ب) $\mu = 80 \pm 1.96(6) = 80 \pm 11.76$

أى أن μ تقع بين 68.24 و 91.76 بمستوى ثقة 95% .

(جـ) $\mu = 80 \pm 2.58(6) = 80 \pm 15.48$

أى أن μ تقع بين 64.52 و 95.48 بمستوى ثقة 99% .

(د) النتائج فى (أ) ، (ب) و (جـ) تشير إلى أنه مع زيادة درجة الثقة المطلوبة ، فإن حجم فترة الثقة يزيد أيضاً ويصبح التقدير بفترة أكثر غموضاً (أى أقل دقة) . ولكن درجة الثقة المرتبطة بفترة ثقة ضيقة جداً قد تكون منخفضة بدرجة تفقد معها معناها . وكتقليد ، فإن فترات الثقة الأكثر استخداما هى 95% ثم 90% و 99% .

٤ - ١٦ أخذت عينة عشوائية من 36 طالباً من بين 500 طالب بمدرسة ثانوية ، متقدمين لامتحان القبول بالجامعة . ووجد أن متوسط درجات العينة هو 380 ، والانحراف المعيارى للمجتمع كله المكون من 500 طالب هو 40 . أوجد فترة الثقة 95% للوسط غير المعلوم للدرجات فى المجتمع كله .

حيث أن $n > 30$ ، فإن توزيع المعاينة النظرى للمتوسط يكون طبيعياً تقريباً . وحيث أن $n > 005\,N$ فإن

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} = \frac{40}{\sqrt{36}} \sqrt{\frac{500-36}{500-1}} \cong \frac{40}{6}(0.96) \cong 6.4$$

ويكون

$$\mu = \bar{X} \pm z\sigma_{\bar{X}} = 380 \pm 1.96(6.4) = 380 \pm 12.54$$

أى أن μ تقع بين 367.46 و 392.54 بمستوى ثقة قدره 95% .

٤ — ١٧ يرغب باحث فى تقدير متوسط الأجر الأسبوعى لعدة آلاف من العاملين بأحد المصانع فى حدود زائد وناقص 20 $ وبدرجة ثقة 99% . ويعرف الباحث من خبرته الماضية أن توزيع الأجر الأسبوعى للعاملين يتبع التوزيع الطبيعى بانحراف معيارى قدره 40 $. ما هو الحد الأدنى للعينة المطلوب ؟

(أ) (بفرض $n < 0.05N$) .

$$z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}}$$

$$z\sigma_{\bar{X}} = \bar{X} - \mu$$

$$z\frac{\sigma}{\sqrt{n}} = \bar{X} - \mu$$

$$2.58\frac{40}{\sqrt{n}} = 20$$

$$2.58\frac{40}{20} = \sqrt{n}$$

(ب) (مقربا للرقم الصحيح الأعلى) . $n = 5.16^2 = 26.63,$ أو 27

٤ — ١٨ (أ) حل المسألة ٤ — ١٧ بإيجاد معادلة n أولا ثم التعويض فيها للحصول على قيمة n . (ب) لماذا يعتبر موضوع حجم العينة مهما ؟ (جـ) ما هو حجم فترة الثقة الإجمالى فى المسألة ٤ — ١٧ ؟ (د) ماذا يكون عليه حجم العينة فى المسألة ٤ — ١٧ لو لم نعرف أن المجتمع يتبع التوزيع الطبيعى ؟ (هـ) ماذا كان يحدث لو لم يكن الانحراف المعيارى للمجتمع معلوماً ؟

(أ) بدءاً بالمقدار $z\sigma/\sqrt{n} = \bar{X} - \mu$ (أنظر المسألة ٤ — ١٧) ، نحصل على $\sqrt{n} = z\sigma/(\bar{X} - \mu)$. فيكون ،

$$n = \left(\frac{z\sigma}{\bar{X} - \mu}\right)^2 \qquad (٤ - ١٠)$$

بالتعويض بالقيم من المسألة ٤ — ١٧ نحصل على

$$n = \left[\frac{(2.58)(40)}{20}\right]^2 = 26.63,$$ أو 27 (نفس النتيجة كما فى المسألة ٤ — ١٧) .

(ب) يعتبر موضوع حجم العينة مهما لأنه إذا كانت العينة صغيرة أكثر من اللازم ، فإننا نفشل فى الوصول إلى أهداف التحليل ، وإذا كانت العينة أكبر مما ينبغى ، فإننا نبدد الموارد لأن التكلفة تكون أعلى عند جمع وتحليل عينة أكبر .

(جـ) حجم فترة الثقة الإجمالى فى المسألة ٤—١٧ هو 40$ ، أى ضعف $\bar{X} - \mu$. وحيث أننا نستخدم $\bar{X}$ كتقدير μ ، فإنه يشار أحياناً إلى $\bar{X} - \mu$ بخطأ التقدير . ولأننا فى المسألة ٤—١٧ نرغب أن يكون خطأ التقدير « فى حدود زائد أو ناقص 20 $ » ، فإن $\bar{X} - \mu = \pm \$20$ ، أى بمدى 40$ لفترة الثقة الإجمالية .

(د) لو لم نعرف أن المجتمع يتبع التوزيع الطبيعى لكان علينا أن نرفع حجم العينة إلى 30 على الأقل فى المسألة ٤ — ١٧ حتى يمكن تبرير استخدام التوزيع الطبيعى .

(هـ) لو لم يكن الانحراف المعيارى σ معلوماً ، لما تمكنا من حل المسألة . (حيث أننا كنا بصدد تحديد حجم العينة الواجب أخذه فى المسألة ٤ — ١٧ ، فإنه لا يمكننا استخدام s كتقدير للانحراف المعيارى σ) . والطريقة الوحيدة لتقدير σ (وبالتالى تقدير قيمة تقريبية لحجم العينة n) تكون إذا عرفنا المدى بين أعلى أجر وأدنى أجر . وحيث أن $\pm 3\sigma$ يتضمن 99.7% من المساحة تحت المنحنى الطبيعى ، كان فى إمكاننا أن نساوى بين 6σ وبين مدى الأجور ومن ثم نقدر σ (وبالتالى نحل المسألة) .

٤ — ١٩ بالإشارة إلى توزيع ذى الحدين ، أذكر العلاقة بين (أ) μ و $\mu_{\bar{p}}$ (ب) p و $\bar{p}$ ، و (جـ) σ، $\sigma_{\bar{p}}$، و $\hat{\sigma}_{\bar{p}}$.

(أ) $\mu = np$ = متوسط عدد النجاحات فى n محاولة ، حيث p احتمال النجاح فى المحاولة الواحدة (أنظر قسم ٣–٣) .
$\mu_{\bar{p}} = \mu/n = p$ وتساوى نسبة النجاحات فى توزيع المعاينة للنسبة .

(ب) p = نسبة النجاحات فى المجتمع و $\bar{p}$ = نسبة النجاحات فى العينة (وهى مقدر غير متحيز للنسبة p) .

(جـ) $\sigma = np(1-p)$ = الانحراف المعيارى لعدد النجاحات فى المجتمع ، و

$$\sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}}$$ = الخطأ المعيارى للنسبة p (٤ – ٦ أ)

أو

$$\sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}}\sqrt{\frac{N-n}{N-1}}$$ عندما $n > 0.05N$ (٤ – ٦ ب)

$$\hat{\sigma}_{\bar{p}} = \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}$$ أو $$\hat{\sigma}_{\bar{p}} = \sqrt{\frac{\bar{p}(1-\bar{p})}{n}}\sqrt{\frac{N-n}{N-1}}$$ عندما $n > 0.05N$ (٤ – ١١ أ ، ب)

٤ — ٢٠ فى عينة عشوائية حجمها 100 عامل من مصنع به 1,200 عامل ، وجد أن 70 يفضلون الاشتراك فى نظام المعاشات كأفراد بدلا من الاشتراك فى مشروع معاشات خاص بالشركة . أوجد فترة الثقة 95% لنسبة العاملين الذين يفضلون مشروعات معاشات فردية .

$$\bar{p} = \frac{70}{100} = 0.7$$

$$p = \bar{p} \pm z\sigma_{\bar{p}}$$ حيث $n > 30$ و $np > 5$ و $n(1-p) > 5$

$$= \bar{p} \pm z\sqrt{\frac{p(1-p)}{n}}\sqrt{\frac{N-n}{N-1}}$$ حيث $n > 0.05N$

$$= 0.7 \pm 1.96\sqrt{\frac{(0.7)(0.3)}{100}}\sqrt{\frac{1,200-100}{1,200-1}}$$ باستخدام $\bar{p}$ كتقدير للنسبة p

$$\cong 0.7 \pm 1.96(0.05)(0.96)$$

$$\cong 0.7 \pm 0.09$$

وعليه فإن p (نسبة كل العاملين فى المصنع الذين يفضلون مشروعات معاشات فردية) تقع بين 0.61 و 0.79 بدرجة ثقة 95% .

٤ — ٢١ ترغب هيئة لاستطلاع الرأى العام أن تقدر بمستوى ثقة 90% نسبة الناخبين المتوقع أن يعطوا أصواتهم لمرشح معين فى حدود 0.06 ± من النسبة الحقيقية (للمجتمع) بين الناخبين . ما الحد الأدنى لحجم العينة إذا كانت استطلاعات أخرى تشير إلى أن نسبة المصوتين لهذا المرشح هى 0.30 ؟

$$z = \frac{\bar{p} - p}{\sigma_p}$$

$$z\sigma_p = \bar{p} - p$$

$$z\sqrt{\frac{p(1-p)}{n}} = \bar{p} - p \qquad \text{بفرض } n < 0.05N$$

$$1.64\sqrt{\frac{(0.3)(0.7)}{n}} = 0.06$$

$$\frac{2.6896(0.3)(0.7)}{n} = 0.0036 \qquad \text{بتربيع الطرفين}$$

$$n = \frac{(2.6896)(0.3)(0.7)}{0.0036} \cong 156.89, \text{ أو } 157$$

٤ – ٢٢ (أ) حل المسألة ٤ –٢١ بإيجاد معادلة n أولا ثم بالتعويض فيها للحصول على قيمة n (ب) كيف كان يمكننا حل المسألة ٤ – ٢١ لو لم نعرف أن نسبة المصوتين للمرشح كانت 0.30؟

(أ) بدءاً بالمقدار $z\sqrt{p(1-p)/n} = \bar{p} - p$ (أنظر المسألة ٤ – ٢١) نحصل على

$$\frac{z^2p(1-p)}{n} = (\bar{p} - p)^2 \qquad \text{و} \qquad n = \frac{z^2p(1-p)}{(\bar{p} - p)^2} \qquad (٤ - ١٢)$$

وبالتعويض بالقيم من المسألة ٤ – ٢١ ، نحصل على

$$n = \frac{(1.64)^2(0.3)(0.7)}{0.06^2} = \frac{(2.6896)(0.21)}{0.0036} \cong 156.89, \text{ أو } 157$$

(نفس الإجابة كما فى المسألة ٤ – ٢١) .

(ب) إذا لم نكن نعرف أن نسبة المصوتين للمرشح كانت 0.30 فيمكن تقدير أكبر قيمة لحجم العينة n للحفاظ على درجة الدقة المطلوبة مهما كانت القيمة الفعلية للنسبة p . ويكون هذا بوضع $p = 0.5$ (فتكون $1 - p = 0.5$ أيضاً) . وحيث أن $p(1-p)$ تظهر فى بسط معادلة n (أنظر أ) فإن المحصلة تكون أكبر ما يكون عندما p و $1 - p$ كلاهما تساوى 0.5 ، وتكون n أكبر ما يمكن ، وعليه

$$n = \frac{z^2p(1-p)}{(\bar{p} - p)^2} = \frac{1.64^2(0.5)(0.5)}{0.06^2} = \frac{(2.6896)(0.25)}{0.0036} \cong 186.8, \text{ أو } 187$$

(بدلا من $n = 157$ عندما علمنا أن $p = 0.30$). وفى هذه الحالة والحالات المشابهة ، فإن محاولة الحصول على تقدير فعلى للنسبة p لا يخفض حجم العينة المطلوب كثيراً . وعندما نفترض p تساوى 0.5 ، فإن معادلة n يمكن تبسيطها إلى

$$n = \left[\frac{z}{2(\bar{p} - p)}\right]^2 \qquad (٤ - ١٣)$$

وباستخدام هذه الأخيرة نحصل على

$$n = \left[\frac{1.64}{2(0.06)}\right]^2 = \left(\frac{1.64}{0.12}\right)^2 \cong 186.8,\ 187$$ أو (نفس القيمة بماليه)

فترات الثقة للوسط باستخدام توزيع t

٤ – ٢٣ (أ) فى أى ظروف لا يمكننا استخدام التوزيع الطبيعى ولكن يمكننا استخدام توزيع t لإيجاد فترات الثقة لوسط المجتمع غير المعلوم ؟ (ب) ما هى العلاقة بين توزيع t والتوزيع الطبيعى القياسى؟ (ج) ما هى العلاقة بين إحصاءات z و t لتوزيع المعاينة النظرى للمتوسط ؟ (د) ماذا يقصد بدرجات الحرية ؟

(أ) عندما يتبع المجتمع التوزيع الطبيعى ولكن σ غير معلومة وحجم العينة ، n ، أصغر من 30 ، فإنه لا يمكننا استخدام التوزيع الطبيعى لتحديد فترات الثقة لوسط المجتمع غير المعلوم ، ولكن يمكننا استخدام توزيع t .

(ب) مثل التوزيع الطبيعى القياسى ، فإن توزيع t هو أيضاً جرسى الشكل ومتماثل حول الوسط الحسابى صفر ولكنه مفرطح (أنظر قسم ٢ – ٤) أو أكثر انبساطاً من التوزيع الطبيعى القياسى ، وبالتالى فجزء أكبر من مساحته يقع عند الأطراف . وبينما هناك توزيع طبيعى قياسى واحد ، فإن هناك توزيع t مختلفاً لكل حجم عينة n . ولكن ، مع تزايد n فإن توزيع t يقترب من التوزيع الطبيعى القياسى حتى يتساويا تقريباً عند $n \geq 30$.

(جـ)

$$z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

ونكشف عنها فى ملحق ٣ .

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}} \qquad (٤ - ١٤)$$

ونكشف عنها فى ملحق ٥ لدرجات الحرية المناظرة .

(د) درجات الحرية (df) تشير إلى عدد القيم التى لنا حرية اختيارها . فمثلا ، إذا كنا نتعامل مع عينة من مفردتين ونعلم أن متوسط العينة لهاتين المفردتين هو 10، فإن لنا حرية تحديد قيمة مفردة واحدة منهما.. فإذا كان الرقم الأول 8 فإن الرقم الثانى يجب أن يكون 12 (لكى يكون المتوسط 10). عندئذ نقول أن لدينا $n-1=2-1=1$ df. وبالمثل ، إذا كانت $n = 10$ ، فإن ذلك يعنى أن لدينا حرية اختيار 9 قيم منها إذا أردنا تقدير متوسط المجتمع ، وعليه يكون لدينا $n - 1 = 10 - 1 = 9$ df .

٤ – ٢٤ (أ) كيف يمكن إيجاد قيمة t التى تناظر 10% من المساحة عند الأطراف وبدرجات حرية 9 ؟ (ب) كيف تفسر قيم t مختلفاً عن تفسير قيم z ؟ (جـ) أوجد قيم t المناظرة لنسب 5 ، 2.5 و 0.5% من المساحة عند الأطراف لعدد 9 من درجات الحرية . (د) أوجد قيم المناظرة لنسب 5 ، 2.5 و 0.5% عند الأطراف لحجم عينة ، n ، كبير جداً أو لا نهائى . كيف تقارن بين t هذه بقيم z المناظرة ؟

(أ) يمكن الحصول على قيمة t المناظرة لنسبة 10% من المساحة عند الأطراف بالتحرك عبر العمود الذى رأسه 0.10 فى ملحق ٥ حتى نصل إلى درجات حرية 9 . وهذا يعطى قيمة t تساوى 1.383 . وبالتماثل ، فإن 10% من المساحة لتوزيع t بدرجات حرية 9 تقع عند الطرف الأيسر ، إلى اليسار من $t = -1.383$.

(ب) تشير قيم t فى ملحق ٥ إلى المساحات (الاحتمالات) عند أطراف توزيع t المقابلة لدرجات الحرية المعينة . أما قيم z فى ملحق ٣ فإنها تشير إلى المساحات (الاحتمالات) تحت المنحنى الطبيعى المعيارى التى تقع بين المتوسط وبين قيم z المحددة (قارن مثال ٤ بمثال ٨) .

(جـ) بالتحرك عبر الأعمدة التى رؤوسها 0.05 ، 0.025 ، و 0.005 فى ملحق ٥ حتى نصل إلى 9 df ، نحصل على قيم t 1.833 ، 2.262 و 3.250 على الترتيب .. وكنتيجة للتماثل فإن 5 ، 2.5 و 0.5% من المساحة تقع فى الطرف الأيسر لتوزيع t لدرجات حرية 9 إلى اليسار من $t = -1.833$ ، $t = -2.262$ ، $t = -3.350$ على الترتيب .

(د) عندما تكون حجوم العينات (ودرجات الحرية) كبيرة جداً أو لا نهائية فإن قيمة $t_{0.05} = 1.645$ ، $t_{0.025} = 1.960$ و $t_{0.005} = 2.576$ (من الصف الأخير فى ملحق ٥) . وهذه تتطابق مع قيم z المناظرة فى ملحق ٣ . بالتحديد $t_{0.025} = 1.960$ تعنى أن 2,5% من المساحة تحت توزيع t بدرجات حرية ∞ تقع عند الطرف الأيمن ، إلى اليمين من $t = 1.96$. وبالمثل ، فإن $z = 1.96$ تعطى (من ملحق ٣) 0.4750 من المساحة تحت التوزيع الطبيعى القياسى من $\mu=0$ إلى $z = 1.96$. وعليه ، لدرجات حرية هى $\text{df} = n-1 = \infty$ فإن توزيع t يتطابق مع التوزيع القياسى الطبيعى .

٤ — ٢٥ أخذت عينة عشوائية مكونة من 25 مفردة بمتوسط 80 وانحراف معيارى 30 من مجتمع مكون من 1,000 مفردة ويتبع التوزيع الطبيعى . أوجد فترات الثقة الآتية لوسط المجتمع غير المعلوم (أ) 90% (ب) 95% (جـ) 99% (د) كيف تقارن هذه النتائج بنتائج المسألة ٤ — ١٥ ؟

(أ) $t_{0.05} = 1.711$ عند 24 df

$$\mu = \bar{X} \pm t\frac{s}{\sqrt{n}} = 80 \pm 1.711\frac{30}{\sqrt{25}} = 80 \pm 1.711\sigma = 80 \pm 10.266$$

أى أن μ تقع بين 69.734 و 90.266 بمستوى ثقة 90%

(ب) $t_{0.025} = 2.064$ عند 24 df

$$\mu = \bar{X} \pm t\frac{s}{\sqrt{n}} = 80 \pm 2.064\frac{30}{\sqrt{25}} = 80 \pm 12.384$$

أى أن μ تقع بين 67.615 و 92.284 بمستوى ثقة 95%

(جـ) $t_{0.005} = 2.797$ عند 24 df

$$\mu = \bar{X} \pm t\frac{s}{\sqrt{n}} = 80 \pm 2.797\frac{30}{\sqrt{25}} = 80 \pm 16.782$$

أى أن μ تقع بين 63.218 و 96.782 بمستوى ثقة 99% .

(د) فترات الثقة 90 ، 95 و 99% كما هو متوقع ، أكبر فى هذه المسألة ، حيث استخدم توزيع t ، عنها فى المسألة ٤—١٥ ، عندما استخدمنا التوزيع الطبيعى القياسى . ولكن الفرق ليس كبيراً لأنه عند $n = 25$ فإن توزيع t

والتوزيع الطبيعى القياسى يتقاربان إلى حد كبير . لاحظ أننا فى هذه المسألة استخدمنا توزيع t لأن المتاح هو s (وليس σ ، كما فى المسألة ٤ – ١٥) .

٤ – ٢٦ سحبت عينة عشوائية مكونة من $n = 9$ مصابيح كهربائية بمتوسط عمر 300 ساعة وانحراف معيارى s قيمته 45 ساعة من شحنة كبيرة من المصابيح الكهربائية معروف أن عمر تشغيلها يتبع التوزيع الطبيعى . (أ) أوجد فترة الثقة 90% لوسط عمر التشغيل غير المعلوم للشحنة كلها . (ب) وضح بالرسم النتائج فى (أ) .

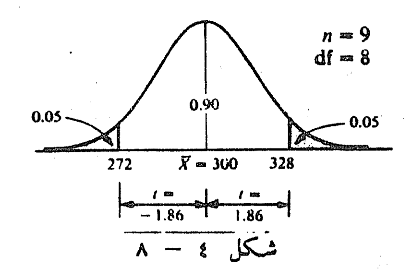
(أ)

$t_{0.05} = 1.860$ عند 8 df

$$\mu = \bar{X} \pm t\frac{s}{\sqrt{n}} = 300 \pm 1.860\frac{45}{\sqrt{9}} = 300 \pm 27.9$$

أى أن μ تقع تقريباً بين 272 و 328 ساعة بمستوى ثقة 90% .

(ب) أنظر شكل ٤ – ٨



شكل ٤ – ٨

٤ – ٢٧ أخذت عينة تشوائية عدد مفرداتها $n = 25$ بمتوسط $\bar{X} = 80$ من مجتمع 1,000 انحرافه المعيارى $\sigma = 30$. افترض أننا نعرف أن المجتمع الذى أخذت منه العينة لا يتبع التوزيع الطبيعى . (أ) أوجد فترة الثقة 95 % لوسط المجتمع غير المعلوم (ب) كيف تقارن هذه النتيجة بالنتائج فى المسألتين ٤ – ١٥ (ب) ، ٤ – ٢٥ (ب) ؟

(أ) حيث أننا نعرف أن المجتمع الذى أخذت منه العينة لا يتبع التوزيع الطبيعى وأن $n < 30$ ، فإننا لا نستطيع استخدام التوزيع الطبيعى أو توزيع t . ولكن يمكننا استخدام نظرية تشبتشيف ، والتى تنص على أنه بصرف النظر عن شكل التوزيع ، فإن نسبة المشاهدات (أو المساحة) التى لا تبعد عن الوسط بأكثر من K انحراف معيارى) هى على الأقل $1 - (1/K^2)$ حيث $K \geq 1$ (أنظر المسألة ٣ – ٤٠) . بوضع $1 - (1/K^2) = 0.95$ والحل لإيجاد K ، نحصل على

$$\frac{1}{K^2} = 1 - 0.95$$

$$1 = 0.05K^2$$

وعليه

$$K^2 = 20$$

$$K \cong 4.47$$

$$\mu = \bar{X} + K\frac{\sigma}{\sqrt{n}} = 80 \pm 4.47\frac{30}{\sqrt{25}} \cong 80 \pm 26.82$$

أى أن μ تقع تقريباً بين 53 و 107 بمستوى ثقة 95% .

(ب) إن فترة الثقة 95 باستخدام نظرية تشبتشيف أوسع كثيراً من تلك السابق إيجادها باستخدام التوزيع الطبيعي (المسألة ٤-١٥ (ب)) ، أو باستخدام توزيع t (المسألة ٤-٢٥ (ب)). ولهذا السبب ، فإنه من النادر استخدام نظرية تشبتشيف لإيجاد فترات الثقة للوسط غير المعلوم للمجتمع . ولكنها تمثل الاختيار الوحيد إذا لم يمكن زيادة حجم العينة إلى 30 على الأقل (حتى يمكن استخدام التوزيع الطبيعي) .

٤ – ٢٨ في أي ظروف يمكن تكوين فترات الثقة للوسط غير المعلوم للمجتمع من عينة عشوائية مسحوبة من المجتمع باستخدام (أ) التوزيع الطبيعي ؟ (ب) توزيع t ؟ (جـ) نظرية تشبتشيف ؟

(أ) يمكن استخدام التوزيع الطبيعي(١) إذا كان المجتمع الأصلي طبيعياً $n \geq 30$ و σ, أو s معلومة (٢) إذا كانت $n \geq 30$ (بالالتجاء إلى نظرية النهاية المركزية) وباستخدام s كتقدير للانحراف المعياري σ ، (٣) إذا كانت $n < 30$ ولكن σ معلومة والمجتمع الذي أخذت منه العينة العشوائية من المعروف أنه يتبع التوزيع الطبيعي .

(ب) يمكن استخدام توزيع t (لدرجات الحرية المعينة) عندما $n < 30$ ولكن σ غير معلومة والمجتمع الذي أخذت منه العينة من المعروف أنه يتبع التوزيع الطبيعي .

(جـ) عندما $n < 30$ و σ غير معلومة والمجتمع الذي أخذت منه العينة لايتبع التوزيع الطبيعي، فن الناحية النظرية لا نستطيع استخدام التوزيع الطبيعي أو توزيع t . في مثل هذه الحالة ، أما أن نستخدم نظرية تشبتشيف وأما أن نرفع من حجم العينة العشوائية إلى $n \geq 30$ (لكي نستطيع استخدام التوزيع الطبيعي) . ومع ذلك فالواقع أن توزيع t يستخدم حتى في هذه الحالات .

مسائل اضافيه

المعـــاينة :

٤ – ٢٩ (أ) ماذا يعني الاستدلال الإحصائي ؟ (ب) ما هي أسماء الخواص الوصفية للمجتمعات والعينات ؟ (جـ) كيف يمكن الحصول على عينات ممثلة ؟

الإجابة : (أ) التقدير واختبار الفروض (ب) المعالم والإحصائيات (جـ) بالمعاينة العشوائية .

٤ – ٣٠ بدءاً بالعمود الثالث والصف العاشر في ملحق ٤ وبالقراءة أفقياً ، كون عينة من 5 من بين 99 عنصرا . (ب) بدءاً بالعمود السابع والصف الأول في ملحق ٤ وبالقراءة رأسياً ، كون عينة من 10 مفردات من بين 400 مفردة .

الإجابة : (أ) 31 ، 13 ، 33 ، 67 ، 68 (ب) 24 ،54 ، 290 ، 218 ، 385 ، 130 ، 24 ، 72 ، 313 ، 387

توزيع المعاينة للوسط :

٤ – ٣١ كيف يمكننا الحصول على توزيع المعاينة النظري للوسط من مجتمع (أ) محدود ؟ (ب) غير محدود ؟

الإجابة : (أ) بأخذ كل العينات الممكنة ذات الحجم n من المجتمع ثم إيجاد متوسط كل عينة (ب) بأخذ (فرضا) عدد لا نهائي من العينات من حجم n من المجتمع اللانهائي ثم إيجاد متوسط كل عينة .

٤ – ٣٢ ما هو (أ) الوسط (ب) الخطأ المعياري لتوزيع المعاينة النظري للوسط ؟

الإجابة : (أ) $\mu_{\bar{X}} = \mu$ حيث μ وسط المجتمع الأصلي (ب) $\sigma_{\bar{X}} = \sigma/\sqrt{n}$ حيث σ هي الانحراف المعياري للمجتمع الأصلي و n حجم العينة ، أما للمجتمعات المحدودة من حجم N حيث $n > 0.05N$ فإن

$$n > 0.05N, \; \sigma_{\bar{X}} = (\sigma/\sqrt{n})\sqrt{(N-n)/(N-1)}$$

٤ – ٣٣ بالنسبة لمجتمع مكون من 1,000 مفردة ، بوسط $\mu = 50$ وانحراف معيارى $\sigma = 10$ ما هو الوسط والخطأ المعيارى لتوزيع المعاينة النظرى للوسط لعينة من حجم (أ) 25 و (ب) 81 ؟

الإجابة : (أ) $\mu_{\bar{X}} = 50$ و $\sigma_{\bar{X}} = 2$ (ب) $\mu_{\bar{X}} = 50$ و $\sigma_{\bar{X}} = 1.07$

٤ – ٣٤ ماهو شكل توزيع المعاينة النظرى للوسط لعينة من حجم (أ) 10 إذا كان المجتمع الأصلى طبيعياً ؟ (ب) 50 إذا كان المجتمع الأصلى غير طبيعى ؟ (جـ) علام بنيت إجابتك فى (ب) ؟

الإجابة : (أ) طبيعى (ب) طبيعى تقريباً (جـ) نظرية النهاية المركزية .

٤ – ٣٥ ما هى الإحصائية (أ) لمتغير عشوائى X ؟ (ب) لتوزيع المعاينة النظرى للمتوسط $\bar{X}$ ؟

الإجابة : (أ) $z = (X - \mu)/\sigma$ (ب) $z = (\bar{X} - \mu)/\sigma_{\bar{X}}$

٤ – ٣٦ ما احتمال أن تقع $\bar{X}$ بين 49 و 50 لعينة عشوائية من 36 مفردة من مجتمع بمتوسط $\mu = 48$ و $\sigma = 12$ ؟

الإجابة : 0.1498 أو 14.98%

٤ – ٣٧ ما احتمال أن يقع متوسط عينة من 144 حسابات مدينين مسحوبة من مجتمع به 2,000 من الحسابات بمتوسط 10,000 $ وانحراف معيارى 4,000 $ بين 9.500 $ و 10,500 $ ؟

الإجابة : 0.8812 أو 88.12%

التقدير باستخدام التوزيع الطبيعى :

٤ – ٣٨ ما هى المقدرات بنقطة غير المتحيزة لكل من μ ، σ و p على الترتيب ؟

الإجابة : $\bar{X}$ ، s (كتعريفها فى المعادلات (٢ – ١٠ ب) و (٢ – ١١ ب) و $\bar{p}$.

٤ – ٣٩ باستخدام التوزيع الطبيعى القياسى أذكر فترات الثقة للوسط μ (أ) 90% (ب) 95% ، (جـ) 99%

الإجابة : (أ) $P(\bar{X} - 1.64\sigma_{\bar{X}} < \mu < \bar{X} + 1.64\sigma_{\bar{X}}) = 0.90$

(ب) $P(\bar{X} - 1.96\sigma_{\bar{X}} < \mu < \bar{X} + 1.96\sigma_{\bar{X}}) = 0.95$

(جـ) $P(\bar{X} - 2.58\sigma_{\bar{X}} < \mu < \bar{X} + 2.58\sigma_{\bar{X}}) = 0.99$

٤ – ٤٠ أخذت عينة عشوائية حجمها 144 بمتوسط 300 وانحراف معيارى 100 من مجتمع به 5,000 مفردة . أوجد التقدير بفترة للوسط μ بحيث تكون ثقتنا 90% أن تلك الفترة تتضمن μ

الإجابة : من 286.34 إلى 313.66

٤ – ٤١ بالنسبة للمسألة ٤ – ٤٠ أوجد فترات الثقة (أ) 95% (ب) 99% (جـ) ماذا تقترح إجابات (أ) ، (ب) ؟

الإجابة : (أ) من 283.67 إلى 316.33 (ب) من 278.51 إلى 321.49 (جـ) كلما زادت درجة الثقة ، كلما اتسعت فترة الثقة .

٤ – ٤٢ أخذت عينة من 400 من بين 100,000 مجند بالجيش فى إحدى السنوات ، ووجد أن متوسط وزن المجند فى العينة هو 170 رطلا والانحراف المعيارى لمجتمع المجندين هو 40 رطلا . أوجد فترة الثقة 90% لمتوسط الوزن فى مجتمع المجندين .

الإجابة : من 166.7 إلى 173.3 رطلا .

٤ — ٤٣ ترغب شركة فى تقدير متوسط عدد ساعات التشغيل لنوع معين من المصابيح الكهربائية فى حدود 10 ساعات تشغيل (زائد أو ناقص) وبدرجة ثقة 95% وتعرف الشركة من المعلومات السابقة عن هذا النوع من المصابيح الكهربائية أن $\sigma = 30h$. ما حجم العينة التى يجب أخذها ؟ الإجابة : 35 .

٤ — ٤٤ (أ) أكتب صيغة n لحل المسألة ٤ — ٤٣ . (ب) ما هو حجم فترة الثقة الإجمالى فى المسألة ٤ — ٤٣ ؟ (جـ) ماذا كان يحدث لو كانت $n < 30$ فى المسألة ٤ — ٤٣ ؟

الإجابة : (أ) $n = [z\sigma/(\bar{X} - \mu)]^2$ (ب) 20 ساعة تشغيل (جـ) كان يجب زيادة حجم العينة إلى 30 لتبرير استخدام التوزيع الطبيعى .

٤ — ٤٥ بالنسبة التوزيع ذى الحدين ، أكتب معادلات (أ) μ و σ (ب) σ_p و $\hat{\sigma}_{\bar{p}}$ عندما $n < 0.05N$ (جـ) $\hat{\sigma}_{\bar{p}}$ عندما $n > 0.05N$

الإجابة : (أ) $\mu = np$ and $\sigma = \sqrt{np(1-p)}$ (ب) $\sigma_p = \sqrt{p(1-p)/n}$ and $\hat{\sigma}_p = \sqrt{\bar{p}(1-\bar{p})/n}$

(جـ) $\hat{\sigma}_p = \sqrt{\bar{p}(1-\bar{p})/n} \times \sqrt{(N-n)/(N-1)}$.

٤ — ٤٦ فى عينة مكونة من 36 طالب دراسات عليا فى الاقتصاد من بين 880 طالباً فى نفس البرنامج وجد أن 8 طلاب يحملون درجة جامعية فى الرياضيات . أوجد النسبة بين كل طلاب الدراسات العليا بالجامعة للطلاب الذين يحملون درجة جامعية فى الرياضيات بدرجة ثقة 90% .

الإجابة : من 0.11 إلى 0.33 .

٤ — ٤٧ يرغب صاحب مصنع مصابيح كهربائية تقدير نسبة المصابيح المعيبة فى حدود $\pm$ 0.1 بدرجة ثقة 95% . ما هو الحد الأدنى لحجم العينة المطلوب ، إذا كانت الخبرة السابقة تشير إلى أن نسبة العيب فى المصابيح الكهربائية المنتجة هى 0.2 ؟

الإجابة : 62 .

٤ — ٤٨ (أ) اكتب معادلة n لحل المسألة ٤ — ٤٧ . (ب) كيف كان يمكننا حل المسألة ٤ — ٤٧ إذا كان المنتج لا يعرف أن $p = 0.2$ ؟

الإجابة : (أ) $n = z^2p(1-p)/(\bar{p}-p)^2$ (ب) بوضع $p = 0.5$ فتكون $n = 97$.

فترات الثقة للوسط باستخدام توزيع t :

٤ — ٤٩ أوجد قيمة t لعدد 29 درجة حرية للمساحات التالية الواقعة فى الطرف الأيمن من توزيع t : (أ) 10% (ب) 5% (جـ) 2.5% (د) 0.05% .

الإجابة : (أ) $t_{0.10} = 1.311$ (ب) $t_{0.05} = 1.699$ (جـ) $t_{0.025} = 2.045$ (د) $t_{0.005} = 2.756$.

٤ — ٥٠ أوجد قيمة z المناظرة للمساحات التالية تحت التوزيع الطبيعى القياسى والواقعة بين الوسط وبين z (أ) 40% (ب) 45% (جـ) 47.5% (د) 49.5% (هـ) كيف تقارن قيم z هذه بقيم t المناظرة السابق إيجادها فى المسألة ٤ — ٤٩ ؟

الإجابة (أ) $z = 1.28$ (ب) $z = 1.65$ (جـ) $z = 1.96$ (د) $z = 2.58$ (هـ) القيم المناظرة لكل من z و t متقاربة جداً (قارن $z = 1.28$ مع $t = 1.311$ ، $z = 1.65$ مع $t = 1.699$ ، $z = 1.96$ مع $t = 2.045$ ، و $z = 2.58$ مع $t = 2.756$) .

٤ – ٥١ أخذت عينة عشوائية حيث $n = 16$ بمتوسط $\overline{X} = 50$ وانحراف معيارى $\sigma = 10$ من مجتمع كبير جداً يتبع التوزيع الطبيعى . (أ) أوجد فترة الثقة 95% للوسط غير المعلوم للمجتمع (ب) كيف تكون الإجابة مختلفة لو أن $\sigma = 10$ ؟

الإجابة : (أ) من 44.67 إلى 55.53 (باستخدام توزيع t بدرجات حرية 15) (ب) من 45.1 إلى 54.9 (باستخدام التوزيع الطبيعى القياسى) .

٤ – ٥٢ فى امتحان إحصاء الفصل كبير ، أخذت عينة عشوائية حيث $n = 4$ طالب فكان متوسط الدرجات $\overline{X} = 75$ والانحراف المعيارى للدرجات $s = 8$ ومن المعروف أن الدرجات فى الفصل كله تتبع التوزيع الطبيعى . (أ) أوجد فترة الثقة 95% و (ب) فترة الثقة 99% لوسط الدرجات غير المعلوم فى المجتمع .

الإجابة : (أ) من 62 إلى 88 تقريباً (ب) من 52 إلى 98 تقريباً .

٤ – ٥٣ أخذت عينة عشوائية حيث $n = 16$ بمتوسط $\overline{X} = 50$ وانحراف معيارى $s = 10$ من مجتمع كبير جداً لا يتبع التوزيع الطبيعى (أ) أوجد فترة الثقة 95% للوسط غير المعلوم للمجتمع (ب) كيف تختلف الإجابة فى (أ) عن تلك فى المسألة ٤ – ٥١ ؟

الإجابة : (أ) من 39 إلى 61 « باستخدام نظرية تشبتشيف واستخدام s كتقدير تقريبى بدلا من σ (ب) فترة الثقة 95% أوسع كثير هنا من تلك فى المسألة ٤ – ٥١ .

٤ – ٥٤ أذكر أى توزيع ينبغى استخدامه لإيجاد فترات الثقة للوسط غير المعلوم للمجتمع من عينة عشوائية مأخوذة من المجتمع فى الحالات التالية (أ) $n = 36$ و $s = 10$ ، (ب) $n = 20$ و $s = 10$ والمجتمع يتبع التوزيع الطبيعى (ج) $n = 20$ و $s = 10$ والمجتمع لا يتبع التوزيع الطبيعى .

الإجابة : (أ) التوزيع الطبيعى (باستخدام نظرية النهاية المركزية واستخدام s كتقدير بدلا من σ) (ب) توزيع t بدرجات حرية 19 (ج) نظرية تشبتشيف .

الفصل الخامس

الاستدلال الاحصائى : اختبار الفروض

٥ــ١ اختبار الفروض

اختبار الفروض عن خصائص المجتمع (مثل μ و σ) هو جانب أساسى آخر من جوانب الاستدلال والتحليل الإحصائى . وفى اختبار الفروض نبدأ بعمل فرض ما عن خاصية المجتمع غير المعلومة . ثم نأخذ عينة عشوائية من المجتمع ، وعلى أساس الخاصية المناظرة فى العينة ، أما أن نقبل وإما أن نرفض الفرض بدرجة ثقة محددة .

وفى اختبار الفروض يمكن أن نرتكب نوعين من الخطأ . الأول ، أنه يمكن أن نرفض على أساس من معلومات العينة فرضاً بينما هو صحيح فى الواقع . ويسمى هذا **خطأ من النوع الأول** . والثانى ،أنه يمكن أن نقبل فرضاً خاطئاً ويسمى هذا **خطأ من النوع الثانى** .

ويمكننا ضبط أو تحديد احتمال ارتكاب خطأ من النوع الأول ، α . ولكن إذا خفضنا α ، فسوف نضطر إلى قبول احتمال أكبر لارتكاب خطأ من النوع الثانى β ، اللهم إلا إذا رفعنا حجم العينة . وتسمى α **مستوى المعنوية** ، و $\alpha - ١$ **مستوى الثقة** للاختبار .

مثال (١) : افترض أن شركة تنتج مصابيح كهربائية ترغب فى معرفة ما إذا كان يمكنها الادعاء بأن مصابيحها الكهربائية تستمر لمدة 1000 ساعة احتراق ، μ . لمعرفة ذلك ، يمكن للشركة أن تأخذ عينة عشوائية من 100 مصباح مثلا وإيجاد متوسط عمرها $\bar{X}$. وكلما صغر الفرق بين $\bar{X}$ و μ ، كلما زادت فرصة قبول الفرض بأن $\mu = 1000$ ساعة احتراق عند مستوى المعنوية المحدد ، α . بوضع α تساوى 5% فإن الشركة تقبل المخاطرة المحسوبة برفض فرض صحيح فى 5% من الحالات . بوضع α عند 1% ، فإن الشركة تواجه باحتمال أكبر لقبول فرض خاطىء ، β .

٥ــ٢ اختبار فروض عن الوسط والنسبة فى المجتمع

الخطوات الرسمية لاختبار فروض عن وسط المجتمع (أو النسبة) هى كالآتى :

١ – افترض أن μ تساوى قيمة افتراضية μ_0 . يمكن تمثيل ذلك بالعبارة $H_0 : \mu = \mu_0$ ويسمى **الفرض العدمى** . وتكون **الفروض البديلة** هى إذن $H_1 : \mu \neq \mu_0$ (وتقرأ « μ لاتساوى μ_0 ») ، $H_1 : \mu > \mu_0$ ، أو $H_1 : \mu < \mu_0$ ، وفقاً للمسألة .

٢ – حدد مستوى معنوية للاختبار (عادة 5 % ، ولكن أحياناً 1 %) وعرف **منطقة القبول** و**منطقة الرفض** . للاختبار باستخدام التوزيع الملائم .

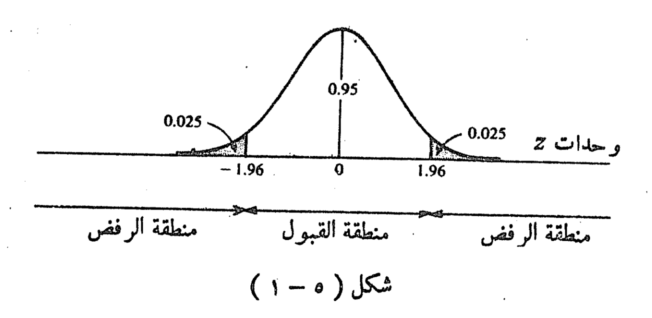
٣ – خذ عينة عشوائية من المجتمع واحسب $\bar{X}$. فإذا وقعت $\bar{X}$ (مقيسة بوحدات الانحراف المعيارى) داخل منطقة القبول ، اقبل H_0 ، وإلا فارفض H_0 لصالح H_1 .

مثال (٢) : افترض أن الشركة فى مثال (١) ترغب فى اختبار ما إذا كان يمكنها الادعاء بأن متوسط عمر المصباح من انتاجها هو 1000 ساعة احتراق . وأنها قامت بأخذ عينة عشوائية حجمها $n = 100$ من إنتاجها فوجدت أن متوسط العينة $\bar{X} = 980$ ساعة والانحراف المعيارى للعينة $s = 80$ ساعة . فإذا أرادت الشركة القيام بالاختبار عند مستوى معنوية 5% ، فعليها أن تمضى كالآتى . حيث أن μ يمكن أن تساوى ، تزيد عن ، أو تقل عن 1,000 ، فإن الشركة يجب أن تضع الفرض العدمى والفرض البديل كالآتى :

$$H_0: \ \mu = 1{,}000 \qquad H_1: \ \mu \neq 1{,}000$$

وحيث أن $n > 30$ ، فإن توزيع المعاينة للوسط يكون تقريباً طبيعياً (ويمكننا استخدام s كتقدير بدلا من σ) . وتكون منطقة القبول للاختبار عند مستوى المعنوية 5 % بين ± 1.96 تحت التوزيع الطبيعى القياسى وتكون منطقة الرفض خارجها (أنظر شكل ٥ — ١) . وحيث أن منطقة الرفض تقع عند ذيلى التوزيع ، فإن الاختبار يسمى اختبار ذو ذيلين . وتكون الخطوة الثالثة إيجاد قيمة المناظرة لقيمة $\bar{X}$:

$$z = \frac{\bar{X} - \mu_0}{\sigma_{\bar{X}}} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{980 - 1{,}000}{80/\sqrt{100}} = \frac{-20}{8} = -2.5$$



شكل (٥ — ١)

وحيث أن قيمة z المحسوبة تقع داخل منطقة الرفض ، فإن على الشركة أن ترفض H_0 أى $\mu = 1{,}000$ وتقبل H_1 أى $\mu \neq 1{,}000$ عند مستوى معنوية 5% .

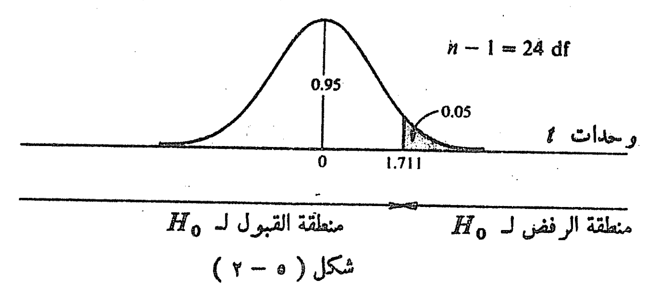
مثال (٣) : ترغب شركة أن تعرف بدرجة ثقة 95 % ما إذا كان يمكنها الادعاء بأن صناديق الصابون المسحوق الذى تبيعه تحتوى على أكثر من 500 جرام (حوالى 1.1 رطل) من الصابون . وتعرف الشركة من الخبرة الماضية أن أوزان الصابون بالصناديق تتبع التوزيع الطبيعى . وقد أخذت الشركة عينة عشوائية حجمها $n = 25$ ووجدت أن $\bar{X} = 520$ جرام و $s = 75$ جرام . وحيث أن الشركة ترغب فى اختبار ما إذا كانت $\mu > 500$ ، فإن

$$H_0:\ \mu = 500 \qquad H_1:\ \mu > 500$$

وحيث أن التوزيع طبيعى ، $n < 30$ ، وكذلك σ غير معلومة ، فعلينا أن نستخدم توزيع t (بدرجات حرية $n - 1 = 24$) لتحديد المنطقة الحرجة ، أى منطقة الرفض ، للاختبار بمستوى معنوية 5% . ونجد ذلك فى ملحق ٥ (أنظر قسم ٤ — ٤) ويعرضها شكل (٥ — ٢) . ويسمى هذا اختبار الذيل الأيمن . وأخيراً ، حيث أن

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}} = \frac{520 - 500}{75/\sqrt{25}} = \frac{20}{15} = 1.33$$

وهى تقع داخل منطقة القبول ، ونقبل H_0 أى $\mu = 500$ ، عند مستوى معنوية 5% (أو بدرجة ثقة 95%) .



شكل (٥ — ٢)

مثال (٤) : تظهر السجلات أن 60 % من الطلاب الذين التحقوا فى الماضى بدراسة جامعية متخصصة قد حصلوا على الدرجة العلمية خلال 4 سنوات . وبالنسبة للملتحقين بالدراسة فى عام ١٩٨٠ وعددهم 36 ، وجد أن 15 طالباً فقط قد حصلوا على الدرجة العلمية حتى ١٩٨٤ . لاختبار ما إذا كانت نتائج الدفعة الملتحقة فى عام ١٩٨٠ أسوأ من نتائج الدفعات السابقة عليها ، فإننا نلاحظ أولا أن المسألة تتعلق بتوزيع ذى الحدين ولكن ، حيث أن $n > 30$ و np و $n(-p) > 5$ ، فإنه يمكننا استخدام التوزيع الطبيعى . (أنظر قسم ٣ — ٥) باستخدام $p = 0.60$ (نسبة النجاح) . بالنسبة لدفعة ١٩٨٠ فإن نسبة النجاح $\bar{p} = 15/36 = 0.42$ والخطأ المعيارى $\sigma_{\bar{p}} = \sqrt{p(1-p)/u} = \sqrt{(0.6)\ (0.4)/36} = 0.08$. وحيث أننا نرغب فى اختبار ما إذا كان أداء دفعة ١٩٨٠ أسوأ من سبقها فإن لدينا

$$H_0:\ p = 0.60 \qquad H_1:\ p < 0.60$$

$$z = \frac{\bar{p} - p}{\sigma_p} = \frac{0.42 - 0.60}{0.08} = -2.25$$

وحيث أن هذا هو **اختبار الذيل الأيسر** وأن 5 % من مساحة التوزيع الطبيعى القياسى تقع إلى اليسار من 1.64 — (أنظر ملحق ٣) ، فإننا نرفض H_0 وننتهى إلى أنه عند مستوى معنوية 5% ، فإن دفعة ١٩٨٠ كانت نتيجتها أسوأ من الدفعات السابقة عليها . ولكن إذا كانت $\alpha = 1\%$ ، فإن المنطقة الحرجة تكون إلى اليسار من $z = -2.33$ وعندئذ نقبل H_0 . وتبين المسألة ٥ — ٥ كيفية تحديد منطقتى القبول والرفض بالوحدات الأصلية للمسألة بدلا من وحدات الانحراف المعيارى . والمسألتان (٥ — ١٠ و ٥ — ١١) تبينان كيفية إيجاد منحنى توصيف العمليات (منحنى OC) ، والذى يعطى قيمة β لقيم μ المختلفة حيث $\mu > \mu_0$. وتبين المسألة (٥ — ١٢) كيفية إيجاد منحنى القوة ، الذى يعطى قيمة α التى تناظر $\mu > \mu_0$.

٥—٣ اختبار الفروض عن الفرق بين وسطين أو الفرق بين نسبتين

فى مواقف اتخاذ قرارات كثيرة ، يكون من المهم تحديد ما إذا كان وسطان أو نسبتان لمجتمعين يتساويان أو يختلفان . ولعمل ذلك فإننا نأخذ عينة عشوائية من كل مجتمع ،فإذا أمكننا أن نعزو الفرق بين وسطى أو نسبتى العينتين إلى الصدفة فإننا ، وفقط فى هذه الحالة ، نقبل فرض أن المجتمعين لهما وسطان (أو نسبتان) متساويان .

إذا كان المجتمعان يتبعان التوزيع الطبيعى (أو إذا كان كل من $n_1, n_2 \geq 30$) فإن توزيع المعـــاينة للفرق بين الوسطين (أو النسبتين) فى العينة يتبع أيضاً التوزيع الطبيعى ، أو يتبع التوزيع الطبيعى تقريباً ، بخطأ معيارى معطى بالمعادلات التالية

(٥ — ١) لاختبار إذا كانت $\mu_1 = \mu_2$

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

(٥ — ٢) لاختبار إذا كانت $p_1 = p_2$

$$\sigma_{\bar{p}_1 - \bar{p}_2} = \sqrt{\frac{\bar{p}(1-\bar{p})}{n_1} + \frac{\bar{p}(1-\bar{p})}{n_2}}$$

حيث (٥ — ٣) متوسط مرجح للنسبتين $\bar{p}_1$ و $\bar{p}_2$

$$\bar{p} = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2}$$

مثال (٥) : ترغب مديرة أن تحدد عند مستوى معنوية 5% ما إذا كان الأجر بالساعة للعمال نصف المهرة متساوياً فى مدينتين . لعمل ذلك ، فإنها تأخذ عينة عشوائية من الأجر بالساعة من كل من المدينتين وتجد أن $\bar{X}_1 = \$ 6.00$ ، $\bar{X}_2 = \$ 5.40$ ، $s_1 = \$ 2.00$ و $s_2 = \$ 1.80$ وذلك لعينتين من حجم $n_1 = 40$ و $n_2 = 54$. الفروض التى يجرى اختبارها هى

$$H_0: \mu_1 = \mu_2 \quad \text{or} \quad H_0: \mu_1 - \mu_2 = 0$$
$$H_1: \mu_1 \neq \mu_2 \quad \text{or} \quad H_1: \mu_1 - \mu_2 \neq 0$$

وهذا اختبار ذو ذيلين وتقع منطقة القبول للفرض H_0 فى حدود ± 1.96 تحت المنحنى الطبيعى القياسى (شكل ٥ – ١)

$$\sigma_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} \cong \sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} = \sqrt{\frac{2.00^2}{40}+\frac{1.80^2}{54}} = \sqrt{0.1+0.06} = \sqrt{0.16} = 0.4$$

$$z = \frac{(\bar{X}_1-\bar{X}_2)-(\mu_1-\mu_2)}{\sigma_{\bar{X}_1-\bar{X}_2}} = \frac{(\bar{X}_1-\bar{X}_2)-0}{\sigma_{\bar{X}_1-\bar{X}_2}} = \frac{0.6}{0.4} = 1.5$$

وحيث أن قيمة z المحسوبة تقع داخل منطقة القبول ، فإننا نقبل H_0 ، أى $\mu_1 = \mu_2$ ، عند مستوى معنوية 5% . ولكن إذا كان من المعروف أن المجتمعين يتبعان التوزيع الطبيعى وكانت كل من n_1 و n_2 أصغر من 30 وافترضنا أن $\sigma_1^2 = \sigma_2^2$ (وكلاهما غير معلوم) ، فإن توزيع المعاينة للفرق بين وسطين يتبع توزيع t بدرجات حرية $n_1 + n_2 - 1$ (أنظر مسألة ٥ – ١٥) .

مثال (٦) : ترغب شركة أن تحدد عند مستوى معنوية 1% ما إذا كانت نسبة المقبول من المكونات الألكترونية لمورد أجنبى ، p_1 تزيد عنها لمورد محلى ، p_2 . وقد أخذت الشركة عينة عشوائية من شحنة كل مورد ووجدت أن $\bar{p}_1 = 0.9$ و $\bar{p}_2 = 0.7$ من عينات من حجم $n_1 = 100$ و $n_2 = 80$. وقد وضعت الشركة الفروض التالية :

$$H_0: p_1 = p_2 \qquad H_1: p_1 > p_2$$

هذا اختبار أيمن الذيل وتقع منطقة الرفض للفرض H_0 إلى اليمين من 2.33 تحت المنحنى الطبيعى القياسى .

$$\bar{p} = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1+n_2} = \frac{(100)(0.9)+(80)(0.7)}{180} = \frac{146}{180} = 0.8$$

$$\sigma_{p_1-p_2} = \sqrt{\frac{\bar{p}(1-\bar{p})}{n_1}+\frac{\bar{p}(1-\bar{p})}{n_2}} = \sqrt{\frac{(0.8)(0.2)}{100}+\frac{(0.8)(0.2)}{80}} = \sqrt{0.0016+0.002} = \sqrt{0.0036} = 0.06$$

حيث

$$z = \frac{(\bar{p}_1-\bar{p}_2)-(p_1-p_2)}{\sigma_{p_1-p_2}} = \frac{0.2}{0.06} = 3.33$$

نرفض H_0 ونقبل الفرض أن $p_1 > p_2$ عند مستوى معنوية 1% .

٥–٤ اختبار كاى – تربيع لجودة التوفيق والاستقلال

يستخدم توزيع كاى – تربيع χ^2 لاختبار (١) إذا كانت التكرارات المشاهدة تختلف « معنويا » عن التكرارات المتوقعة عندما يكون عدد النواتج الممكنة أكثر من اثنين ؛ (٢) إذا كان التوزيع الذى أخذت منه العينة ذا الحدين ، أو الطبيعى ، أو أى توزيع آخر ؛ (٣) إذا كان متغيران مستقلين أم لا .

وإحصائية χ^2 المحسوبة من بيانات العينة معطاة بالصيغة

$$\chi^2 = \sum \frac{(f_0 - f_e)^2}{f_e} \qquad (٥ - ٤)$$

حيث $f_0 =$ التكرارات المشاهدة

$f_e =$ التكرارات المتوقعة

فإذا كانت χ^2 المحسوبة أكبر من القيمة الجدولية χ^2 عند مستوى المعنوية ودرجات الحرية المحددة (من ملحق ٦) ، يرفض الفرض العدمى H_0 ، لصالح الفرض البديل H_1 .

درجات الحرية لاختبار جودة التوفيق (١ و ٢) معطاة بالصيغة

$$df = c - m - 1 \qquad (٥ - ٥)$$

حيث c = عدد الفئات

m = عدد معالم المجتمع التى يجرى تقديرها من إحصائيات العينة .

درجات الحرية لاختبارات الاستقلال أو اختبارات جداول الاقتران (٣) ، معطاة بالصيغة

$$df = (r - 1)(c - 1) \qquad (٥ - ٦)$$

حيث r = عدد الصفوف فى جدول الاقتران

n = عدد الأعمدة

ويكون التكرار المتوقع فى كل خلية من جدول الاقتران

$$f_e = \frac{\sum r \sum c}{n} \qquad (٥ - ٧)$$

حيث n = حجم العينة الإجمالى .

مثال (٧) : وجد محل تجارى من خبرته الماضية أن 30% من التليفزيونات المباعة من الحجم الصغير ، 40% من الحجم المتوسط ، 30% من الحجم الكبير . لتحديد حجم المخزون الواجب الاحتفاظ به من كل نوع ، أخذ المدير عينة عشوائية من 100 من المبيعات الحديثة للتليفزيون فوجد أن منها 20 من النوع الصغير ، 40 من النوع المتوسط ، 40 من النوع الكبير . باستخدام مستوى معنوية 5% ، يختبر المدير الفرض أن نمط المبيعات الماضى لازال سائداً ، ويمضى كالآتى (أنظر جدول ٥ - ١) :

$$\chi^2 = \sum \frac{(f_0 - f_e)^2}{f_e} = \frac{(20-30)^2}{30} + \frac{(40-40)^2}{40} + \frac{(40-30)^2}{30} = \frac{-10^2}{30} + \frac{0^2}{40} + \frac{10^2}{40} = \frac{100}{30} + \frac{100}{40} \cong 5.83$$

$$df = c - m - 1 = 3 - 0 - 1 = 2$$

وحيث أنه لم يتم حساب أى من معالم المجتمع من البيانات فإن $m = 0$. $df = 2$ تعنى أننا إذا علمنا فئتين من الثلاث والمجموع ، فإن الفئة الثالثة لاتكون « حرة » التغير . وحيث أن القيمة المحسوبة $\chi^2 = 5.83$ أصغر من القيمة الجدولية $\chi^2 = 5.99$ بمستوى معنوية $\alpha = 0.05$ ودرجات حرية 2 (أنظر ملحق ٦) ، فإننا لانستطيع أن نرفض H_0 ، بأن نمط المبيعات فى الماضى ما زال سائداً . وعندما يكون التكرار المتوقع فى أى فئة أقل من 5 فإنه يجب ضمها لفئة مجاورة (أنظر المسألة ٥ - ١٨) . لاختبار إذا كان التوزيع موضع المعاينة هو ذا الحدين أو الطبيعى ، أنظر المسألتين (٥ - ١٩) ، (٥ - ٢٠) .

جدول (٥ - ١) المشتريات المشاهدة والمتوقعة لأجهزة التليفزيون حسب حجم الشاشة

	حجم الشاشة			الإجمالى
	كبير	متوسط	صغير	
النمط المشاهد f_0	20	40	40	100
النمط فى الماضى f_e	30	40	30	100

مثال (٨) : جمع تاجر سيارات البيانات الموضحة فى جدول (٥ — ٢) عن عدد السيارات الأجنبية والمحلية التى يشتريها عملاء أعمارهم تحت سن 30 سنة ، والتى يشتريها عملاء أعمارهم سن 30 سنة فأكثر . لاختبار ما إذا كان نوع السيارة المشتراة (أجنبية أو محلية) مستقبلا عن سن المشترى عند معنوية % 1 ، ننشىء جدول التكرارات المتوقعة (جدول ٥ — ٣) . القيمة فى الخلية الأولى صف 1 وعمود 1 ،

$$f_e = \frac{\sum r \sum c}{n} = \frac{(70)(50)}{170} \cong 21$$

ويمكن الحصول على التكرارات المتوقعة الثلاثة الباقية بالطرح من مجموع الصفوف ومجموع الأعمدة . أى

$$\text{df} = (r-1)(c-1) = (2-1)(2-1) = 1$$

$$\chi^2 = \sum \frac{(f_0 - f_e)^2}{f_e} = \frac{(30-21)^2}{21} + \frac{(40-49)^2}{49} + \frac{(20-29)^2}{29} + \frac{(80-71)^2}{71} = 9.44$$

وحيث أن قيمة χ^2 المحسوبة تتجاوز قيمة χ^2 عند $\alpha = 0.01$ و $\text{df} = 1$ (أنظر ملحق ٦) ، نرفض H_0 القائل بأن السن ليس عاملا فى تحديد نوع السيارة المشتراة (وننتهى إلى أن الأصغر سناً يميلون فيما يبدو إلى شراء السيارات الأجنبية . عندما $\text{df} = 1$ ولكن $n < 50$ يستخدم معامل تصحيح للاتصال باستخدام $(|f_0 - f_e| - 0.5)^2$ فى بسط معادلة (٥ — ٤) أنظر المسألة ٥ — ٢٢) .

جدول (٥ — ٢) جدول الاقتران لمشترى السيارات

السن	نوع السيارة		الإجمالى
	محلية	أجنبية	
تحت 30 ،	30	40	70
30 ، فأكثر	20	80	100
إجمالى	50	120	170

جدول (٥ — ٣) جدول التكرارات المتوقعة المناظرة للتكرارات المشاهدة فى جدول (٥ — ٢)

السن	نوع السيارة		الإجمالى
	محلية	أجنبية	
تحت 30 ،	21	49	70
30 فأكثر	29	71	100
إجمالى	50	120	170

٥—٥ تحليل التباين

يستخدم تحليل التباين لاختبار فرض أن متوسطات أكثر من مجتمعين متساوية أو مختلفة عندما تكون المجتمعات موزعة توزيعاً طبيعياً مع تساوى التباين . الخطوات كالآتى :

خطوة (١) : قدر تباين المجتمع من التباين بين متوسطات العينات (MSA فى جدول ٥ — ٤)

خطوة (٢) : قدر تباين المجتمع من التباين داخل العينات (MSE فى جدول ٥ — ٤)

خطوة (٣) : احسب النسبة F (MSA/MSE) فى جدول ٥ — ٤) :

$$F = \frac{\text{التباين بين متوسطات العينات}}{\text{التباين داخل العينات}}$$

خطوة (٤) : إذا كانت F المحسوبة أكبر من قيمة F الجدولية عند مستوى المعنوية ودرجات الحرية المعينة (من ملحق ٧) ، فإن الفرض العدمى H_0 عن تساوى متوسطات المجتمعات ، يرفض لصالح الفرض البديل ، H_1 . الخطوات السابقة موضحة بجدول (٥ — ٤)

جدول (٥ — ٤) جدول تحليل التباين

مصدر التفسير	مجموع المربعات	درجات الحرية	متوسط المربعات	النسبة F
بين الأوساط (يفسره العامل A)	$SSA = r\sum(\bar{X}_J - \bar{\bar{X}})^2$	$c-1$	$MSA = \frac{SSA}{c-1}$	$\frac{MSA}{MSE}$
داخل العينات (الخطأ أو غير المفسر)	$SSE = \sum\sum(X_{iJ} - \bar{X}_J)^2$	$(r-1)c$	$MSE = \frac{SSE}{(r-1)c}$	—
الإجمالى	$SST = \sum\sum(X_{iJ} - \bar{\bar{X}})^2 = SSA + SSE$	$rc-1$	—	—

(٥ — ٨) حيث $\bar{X}_J$ = متوسط العينة J المكونة من r مشاهدة = $(\sum_i X_{iJ})/r$

(٥ — ٩) $\bar{\bar{X}}$ = المتوسط الكبير لكل العينات $(\sum_i\sum_J X_{iJ})/rc$

(٥ — ١٠) SSA = مجموع المربعات التى يفسرها العامل $A = r\sum(\bar{X}_J - \bar{\bar{X}})^2$

(٥ — ١١) SSE = مجموع مربعات الخطأ والتى لا يفسرها العامل $A = \sum\sum(X_{iJ} - \bar{X}_J)^2$

(٥ — ١٢) SST = مجموع المربعات الاجمالى = SSA + SSE = $\sum\sum(X_{iJ} - \bar{\bar{X}})^2$

ويعطى ملحق ٧ قيم F عندما α — 0.05 (الرقم الأعلى) وعندما $\alpha = 0.01$ (الرقم الأسفل) لكل زوج من درجات الحرية :

(٥ — ١٣) درجات حرية البسط = $c - 1$

حيث c = عدد العينات

(٥ — ١٤) درجات حرية المقام = $(r-1)c$

حيث r = عدد المشاهدات فى كل عينة .

مثال (٩) : تبيع شركة نفس الصابون فى ثلاثة أغلفة مختلفة وبنفس السعر . يبين جدول (٥ — ٥) مبيعات 5 شهور . المبيعات موزعة توزيعاً طبيعياً ولها تباين متساو .

جدول (٥ — ٥) مبيعات خمسة شهور من الصابون فى الأغلفة ١ ، ٢ ، ٣

غلاف (٣)	غلاف (٢)	غلاف (١)
90	78	87
91	81	83
84	79	79
82	82	81
88	80	80
435	400	410

لاختبار ما إذا كان متوسط المبيعات لكل غلاف متساوياً أم لا عند مستوى معنوية 5 % (أى $H_0 : \mu_1 = \mu_2 = \mu_3$ مقابل $H_1 : \mu_1, \mu_2, \mu_3$ ليست متساوية) ، تمضى الشركة كالتالى :

$$\bar{X}_1 = \frac{410}{5} = 82 \qquad \bar{X}_2 = \frac{400}{5} = 80 \qquad \bar{X}_3 = \frac{435}{5} = 87$$

$$\bar{\bar{X}} = \frac{410 + 400 + 435}{(5)(3)} = 83$$

$$SSA = 5\left[(82-83)^2 + (80-83)^2 + (87-83)^2\right] = 130$$

$$SSE = (87-82)^2 + (83-82)^2 + (79-82)^2 + (81-82)^2 + (80-82)^2 + (78-80)^2 + (81-80)^2 + (79-80)^2$$
$$+ (82-80)^2 + (80-80)^2 + (90-87)^2 + (91-87)^2 + (84-87)^2 + (82-87)^2 + (88-87)^2$$
$$= 110$$

$$SST = (87-83)^2 + (83-83)^2 + \cdots + (88-83)^2 = SSA + SSE = 240$$

وتستخدم البيانات السابقة لتكوين جدول (٥ — ٦) لتحليل التباين ANOVA

جدول (٥ — ٦) جدول ANOVA لأغلفة الصابون

التغير	مجموع المربعات	درجات الحرية	متوسط المربعات	نسبة
$MSA/MSE = 65/9.17 = 7.09$	$MSA = 130/2 = 65$	$c - 1 = 2$	$SSA = 130$	تفسره الأغلفة (بين الأعمدة)
	$MSE = 110/12 = 9.17$	$(r-1)c = 12$	$SSE = 110$	الخطأ أو غير المفسر (داخل الأعمدة)
	—	$rc - 1 = 14$	$SST = 240$	الإجمالى

وحيث أن القيمة المحسوبة $F = 7.09$ (من جدول ٥ — ٦) تتجاوز القيمة الجدولية $F = 3.88$ عند $\alpha = 0.05$ ودرجات حرية 2 و 12 (أنظر ملحق ٧) فإننا نرفض H_0 ، أى الفرض القائل بأن متوسط المبيعات للأغلفة المختلفة يتساوى ، ونقبل H_1 ، بأنها تختلف . ويشار إلى الإجراء السابق بأنه تحليل التباين فى اتجاه واحد أو لعامل واحد . بالنسبة لتحليل التباين فى اتجاهين أنظر المسألتين (٥ — ٢٦) و (٥ — ٢٧) .

مسائل محلولة

اختبار الفروض :

٥ — ١ (أ) ماذا يقصد باختبار الفروض ؟ ماهو الإجراء العام ؟ (ب) ماذا يقصد بالخطأ من النوع الأول والخطأ من النوع الثانى ؟ (جـ) ماذا يقصد بمستوى المعنوية ؟ بمستوى الثقة ؟

(أ) يشير اختبار الفروض إلى قبول أو رفض ما عن خاصية غير معلومة للمجتمع مثل أحد المعالم أو شكل توزيع المجتمع والخطوة الأولى فى اختبار الفروض هى وضع فرض ما عن خاصية المجتمع غير المعلومة . ثم تؤخذ عينة عشوائية من المجتمع ، وعلى أساس من خاصية العينة المناظرة ، نقبل أو نرفض الفرض بدرجة معينة من الثقة .

(ب) يشير الخطأ من النوع الأول إلى رفض فرض صحيح . ويشير الخطأ من النوع الثانى إلى قبول فرض خاطىء . وفى التحليل الإحصائى ، يمكننا ضبط أو تحديد احتمال الخطأ من النوع الأول أو النوع الثانى . وعادة نعبر عن احتمال الخطأ من النوع الأول أو النوع الثانى . وعادة نعبر عن احتمال الخطأ من النوع الأول بالحرف اليونانى الفا (α) ، بينما نعبر عن احتمال الخطأ من النوع الثانى بالحرف بيتا (β) . وتصغير الخطأ من النوع الأول يترتب عليه زيادة الخطأ من النوع الثانى . والطريقة الوحيدة لتخفيض كل من α و β هو زيادة حجم العينة .

(ﺟ) يشير مستوى المعنوية : إلى احتمال رفض فرض صحيح أى ارتكاب خطأ من النوع الأول (α) . ويشير مستوى الثقة ($1 - \alpha$) إلى احتمال قبول فرض صحيح . وفى العمل الإحصائى ، فإن مستوى المعنوية α ، يحدد عادة عند 5% فيكون مستوى الثقة ، $1 - \alpha$ عند 95% . أحياناً تكون $\alpha = 1\%$ (فتكون $1 - \alpha = 99\%$) .

٥ – ٢ (أ) كيف يمكن اختبار الفرض أن عملة ما متوازنة ؟ (ب) مامعنى كل من الخطأ من النوع الأول والخطأ من النوع الثانى فى هذه الحالة ؟

(أ) لاختبار فرض أن عملة ما متوازنة ، يمكننا رمى العملة عدة مرات وتسجيل عدد مرات الصورة والكتابة . فمثلا ، [illegible] رمى العملة 20 مرة والحصول على 9 صورة بدلا من 10 كالمتوقع ولكن لايعنى هذا بالضرورة أن العملة غير متوازنة . بالتأكيد ، حيث أن 9 « قريبة جداً » من 10 ، « فالمرجح » أننا نتعامل مع عملة متوازنة . ولكن إذا حصلنا فقط على 4 صورة فى 20 رمية ، فنحن على الأرجح نتعامل مع عملة غير متوازنة لأن احتمال الحصول على 4 صورة ، (16 كتابة) فى 20 رمية لعملة متوازنة بالتأكيد صغير جداً (أنظر قسم ٣ – ٣) .

(ب) بالرغم من أن 9 صورة فى 20 رمية يشير على الأرجح إلى عملة متوازنة ، إلا أن هناك دائماً احتمالا صغيراً أن العملة غير متوازنة . وبقبول فرض أن العملة متوازنة ، يمكن أن نكون مرتكبين خطأ من النوع الأول . لكن ، فى حالة 4 صورة فى 20 رمية فإن الأرجح أن العملة غير متوازنة . ولكن بقبول فرض أن العملة غير متوازنة ، فإننا نواجه الاحتمال الصغير بأن العملة متوازنة ، بما يعنى ارتكاب خطأ من النوع الثانى . عند اختبار فرض ما ، يمكن للباحث اختبار احتمال رفض فرض صحيح ، α صغير للدرجة التى يرغبها . ولكن بزيادة « منطقة القبول » للفرض ، على الباحث أن يتقبل بالضرورة احتمال قبول فرض خاطىء أو ارتكاب خطأ من النوع الثانى β .

٥ – ٣ كيف يمكن لمنتج كابلات من الصلب أن يختبر ما إذا كان متوسط مقاومة الكسر للكابلات المنتجة (أ) 5,000 lb (ب) أكبر من 5,000 lb ؟ (ﺟ) أقل من 5,000 lb ؟

(أ) يمكن للمنتج أن يختبر ما إذا كان متوسط قوة المقاومة للكسر للكابلات المنتجة 5,000 lb بأخذ عينة عشوائية من الكابلات وإيجاد متوسط قوة المقاومة للكسر لها ، $\bar{X}$ ، وكلما قربت $\bar{X}$ من القيمة المفترضة $\mu = 5.000$ lb ، كلما كان فى الإمكان أن يقبل المنتج الفرض عند مستوى المعنوية المعين ، α .

(ب) قد يهتم المنتج باختبار ما إذا كان متوسط مقاومة قوة الكسر أكبر من 5,000 lb (أى $\mu > 5,000$ lb) . لعمل ذلك ، مرة أخرى ، يأخذ المنتج عينة عشوائية من الكابلات المنتجة ويختبر متوسط قوة المقاومة للكسر $\bar{X}$. وكلما زادت $\bar{X}$ عن القيمة المفترضة $\mu = 5,000$) كلما كان من الأرجح أن يقبل المنتج الفرض عند مستوى المعنوية المعين ، α

(ﺟ) لاختبار أن متوسط قوة المقاومة للكسر لايتجاوز 5,000 lb ، يوجد المنتج متوسط قوة المقاومة للكسر من عينة عشوائية من كابلات الصلب . وكلما صغرت $\bar{X}$ عن 5,000 lb كلما كان من الأرجح أن يقبل المنتج فرض أن متوسط قوة المقاومة للكسر أقل من 5,000 lb (أى $\mu < 5,000$) ، بدرجة الثقة المحددة ، $1 - \alpha$.

اختبار فروض عن الوسط والنسبة فى المجتمع :

٥ – ٤ يرغب منتج كابلات من الصلب اختبار ما إذا كانت الكابلات التى ينتجها لديها قوة مقاومة للكسر قدرها 5,000 lb . فقوة مقاومة للكسر أقل من 5,000 lb لن تكون ملائمة ، وقوة مقاومة للكسر أكبر من 5,000 lb ترفع التكاليف بدون

مبرر . يأخذ المنتج عينة عشوائية من 64 قطعة ويجد أن متوسط قوة المقاومة للكسر هو 5,100 lb والانحراف المعيارى هو 480 lb . هل يجب أن يقبل المنتج الفرض أن الكابلات الصلب لها قوة مقاومة للكسر 5,000 lb عند مستوى معنوية 5% ؟

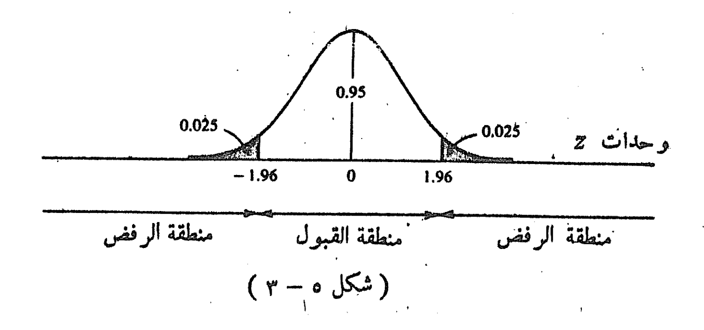
حيث أن μ من الممكن أن تساوى ، تزيد عن ، أو تقل عن 5, 000 lb فإننا نضع الفرض العدمى والفرض البديل كالآتى :

$$H_0: \quad \mu = 5{,}000 \text{ lb} \qquad H_1: \quad \mu \neq 5{,}000 \text{ lb}$$

وحيث أن $n > 30$ ، فإن توزيع المعاينة للوسط طبيعى تقريباً (ويمكن استخدام s كتقدير بدلا من σ) . وتكون منطقة القبول للاختبار عند مستوى معنوية 5% بين 1.96 ± تحت المنحنى الطبيعى القياسى ومنطقة الرفض أو المنطقة الحرجة تكون خارج هذه الحدود (أنظر شكل ٥ – ٣) . وحيث أن منطقة الرفض تقع عند الذيلين فإننا بصدد اختبار له ذيلان . وتكون الخطوة الثالثة إيجاد قيمة z المناظرة لقيمة $\bar{X}$:

$$z = \frac{\bar{X} - \mu_0}{\sigma_{\bar{X}}} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{5{,}100 - 5{,}000}{480/\sqrt{64}} = \frac{100}{60} = 1.67$$

وحيث أن القيمة المحسوبة z تقع داخل منطقة القبول،فيجب أن يقبل المنتج الفرض العدمى H_0 ،ويرفض H_1 عند مستوى معنوية 5% (بمستوى ثقة 95%) . لاحظ أن هذا لا يبرهن أن μ هى بالتأكيد تساوى 5,000 lb ولكنه « يبرهن » فقط على أنه لايوجد شاهد إحصائى على أن μ لاتساوى 5,000 lb عند مستوى المعنوية 5% .



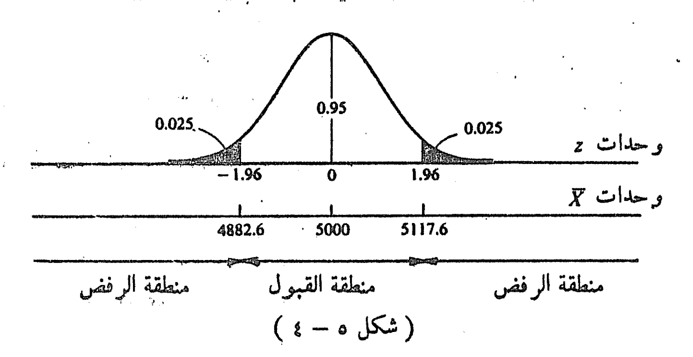
(شكل ٥ – ٣)

٥ – ٥ حدد منطقتى القبول والرفض للمسألة (٥ – ٤) بوحدة الرطل .

لإيجاد منطقة القبول (عند مستوى معنوية 5%) بالرطل ، فإننا نمضى على نمط قسم (٤ – ٤) بإيجاد فترة الثقة 95 % حول μ_0 :

$$\mu_0 \pm z\sigma_{\bar{X}} = \mu_0 \pm z\frac{\sigma}{\sqrt{n}} = \mu_0 \pm z\frac{s}{\sqrt{n}} = 5{,}000 \pm 1.96\frac{480}{\sqrt{64}} = 5{,}000 \pm 117.6$$

أى أنه لقبول H_0 عند مستوى معنوية 5% ، فإن $\bar{X}$ يجب أن تكون أكبر من 4.882. 4 lb وأقل من 5,117.6 lb والعلاقة بين هذا والنتيجة السابق الحصول عليها فى المسألة (٥ – ٤) موضحة فى (شكل ٥ – ٤)



(شكل ٥ – ٤)

٥ – ٦ يعرف مركز تجنيد بالجيش من الخبرة الماضية أن وزن المجند يتبع التوزيع الطبيعى بوسط μ يساوى 80 كيلوجراماً (حوالى 176 رطلا جراما (وانحراف معيارى σ يساوى 10 كيلوجراماً . ويرغب مركز التجنيد أن يختبر، عند مستوى معنوية 1% ، ما إذا كان متوسط وزن مجندى هذا العام أكبر من 80 كيلوجراماً . ولعمل هذا ، فقد أخذ عينة عشوائية من 25 مجنداً حيث وجد أن متوسط الوزن فى العينة 85 كيلوجراماً . كيف يمكن إجراء هذا الأختبار ؟

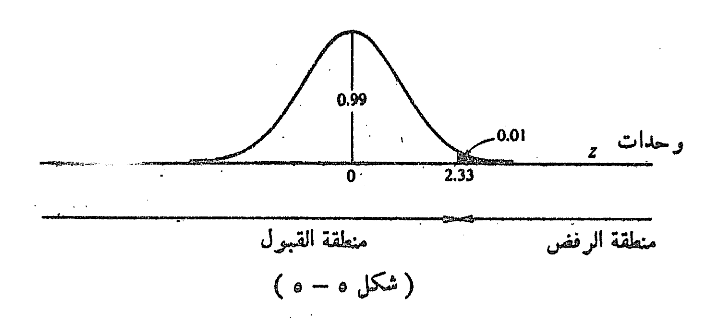
حيث أن المركز يرغب فى اختبار ما إذا كان $\mu > 80$ ، فإنه يضع الفرضين التاليين :

$$H_0: \ \mu = 80 \text{ kg} \qquad H_1: \ \mu > 80 \text{ kg}$$

(تضع بعض الكتب الفرض العدمى كالآتى $H_0 : \mu \leq 80$ ، ولكن النتيجة واحدة) . وحيث أن المجتمع الأصلى يتبع التوزيع الطبيعى وكذلك σ معلومة ، فيمكن استخدام التوزيع الطبيعى القياسى لتحديد المنطقة الحرجة ، أو منطقة الرفض ، للاختبار . وحيث أن $H_1 : \mu > 80$ فإننا بصدد **اختبار الذيل الأيمن** حيث تقع المنطقة الحرجة إلى اليمين من $z = 2.33$ عند مستوى معنوية 1% (أنظر ملحق ٣ وشكل ٥ – ٥) . وعليه فإن

$$z = \frac{\bar{X} - \mu_0}{\sigma_{\bar{X}}} = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{85 - 80}{10/\sqrt{25}} = 2.5$$

وحيث أن قيمة z المحسوبة تقع داخل منطقة الرفض ، فإننا نرفض H_0 ونقبل H_1 (أى $\mu > 80$ kg) . ويعنى هذا أنه إذا كانت $\mu = 80$ kg فإن احتمال أن عينة عشوائية مأخوذة من هذا المجتمع تعطى متوسطاً $\bar{X} = 85$ أقل من 1% . ومثل هذه العينة تكون بالتأكيد غير عادية . وعليه فإننا نرفض H_0 عند مستوى معنوية 1% (أى أننا واثقون 99% من اتخاذ القرار السليم) .



(شكل ٥ – ٥)

٥ – ٧ تتلقى وكالة حكومية شكاوى كثيرة من المستهلكين فحواها أن صناديق مسحوق الصابون التى تبيعها إحدى الشركات تحتوى على كمية أقل من 20 oz من المسحوق المعلن عنه . للتحقق من شكاوى المستهلكين ، اشترت الوكالة 9 صناديق من

المسحوق ووجدت أن $\overline{X} = 18$ oz و $s = 3$oz . كيف يمكن للوكالة إجراء الاختبار عند مستوى معنوية 5 % إذا علم أن كمية المسحوق فى الصناديق موزعة توزيعاً طبيعياً ؟

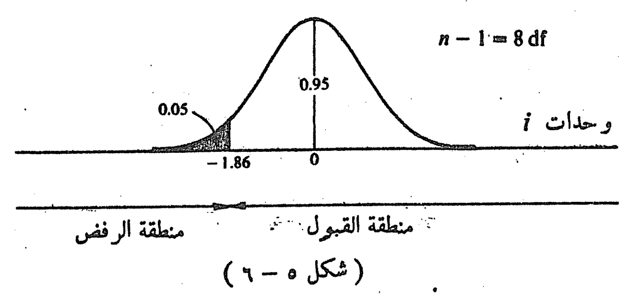
تستطيع الوكالة أن تضع H_0 و H_1 كالتالى :

$$H_0: \mu = 20 \text{ oz} \qquad H_1: \mu < 20 \text{ oz}$$

(تضع بعض الكتب الفرض العدمى كالآتى $H_0 : \mu \geq 20$ ، ولكن النتيجة واحدة) . وحيث أن المجتمع الأصلى يتبع التوزيع الطبيعى ، ولكن σ غير معلومة و $n < 30$ ، فإنه يجب استخدام توزيع t (بدرجات حرية 8 و $s = \sigma$) لتحديد منطقة الرفض لاختبار الذيل الأيسر هذا عند مستوى معنوية 5% (أنظر شكل ٥ – ٦) . فيكون

$$t = \frac{\overline{X} - \mu_0}{\sigma_{\overline{X}}} = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{\overline{X} - \mu_0}{s/\sqrt{n}} = \frac{18 - 20}{3/\sqrt{9}} = -2.0$$

وحيث أن قيمة t المحسوبة تقع داخل منطقة الرفض ، فإن على الوكالة أن ترفض H_0 وتقبل شكاوى المستهلكين ، H_1 . لاحظ أنه لو كانت α قيمتها 1% لوقعت منطقة الرفض إلى اليسار من $t = -2.896$ ، مؤدياً ذلك إلى قبول H_0 . ومن ثم فإنه من المهم تحديد مستوى المعنوية قبل الاختبار .



(شكل ٥ – ٦)

٥ – ٨ يريد مستشفى أن يختبر أن 90% من جرعات عقار يشتريه يحتوى على 100 mg (1/1000 g) من العقار . لعمل هذا ، يأخذ المستشفى عينة من $n = 100$ جرعة ، ويجد أن 95 منها فقط تحتوى على الكمية المناسبة . كيف يمكن المستشفى أن يختبر هذا عند : (أ) $\alpha = 1\%$ ؟ (ب) $\alpha = 5\%$ ؟ (جـ) $\alpha = 10\%$

(أ) تتعلق هذه المشكلة بتوزيع ذى الحدين . ولكن ، طالما أن $n > 30$ و np وكذلك $n(1-p)$ أكبر من 5 ، فيمكن استخدام التوزيع الطبيعى مع $p = 0.90$. بالنسبة للعينة

$$\bar{p} = \frac{85}{100} = 0.85 \qquad \text{and} \qquad \sigma_p = \sqrt{\frac{p(1-p)}{100}} = \sqrt{\frac{(0.9)(0.1)}{100}} = 0.03$$

وحيث أننا نرغب فى إيجاد ما إذا كانت $p \gtrless 0.09$ ، فإن $H_0 : p = 0.90$ و $H_1 : p \neq 0.90$ وتقع منطقة القبول للفرض H_0 عند مستوى معنوية 1 % فى حدود ± 2.58 انحراف معيارى (أنظر ملحق ٣) . وحيث أن

$$z = \frac{\bar{p} - p}{\sigma_p} = \frac{0.85 - 0.90}{0.03} = 1.67$$

فإن على المستشفى أن يقبل H_0 ، أى $p = 0.90$ عند مستوى المعنوية 1% .

(ب) عند مستوى المعنوية 5% ، تقع منطقة القبول المرض H_0 فى حدود 1.96 ± انحراف معيارى ، وعليه فإن المستشفى يجب أيضاً أن يقبل H_0 ويرفض H_1 بدرجة ثقة 95% .

(جـ) عند مستوى المعنوية 10% ، تقع منطقة القبول للفرض H_0 فى حدود 1.64 ± انحراف معيارى (أنظر ملحق ٣) وعليه فإن على المستشفى أن يرفض H_0 ويقبل H_1 ، أى $p \neq 0.90$. لاحظ أن القيم الأعلى للاحصائية α توسع منطقة الرفض H_0 (أى تزيد من احتمال قبول H_1) . علاوة على أنه مع تزايد α (أى تزايد احتمال رفض المرض H_0 بينما هو صحيح (، تتناقص β (احتمال قبول فرض خاطىء) .

٥ – ٩ يدعى متحدث حكومى لمكافحة التلوث أن أكثر من 80% من المصانع فى المنطقة تستوفى معايير مكافحة التلوث . ولكن واحدة من أنصار مكافحة التلوث لاتصدق ادعاء الحكومة . فهى تأخذ عينة عشوائية من البيانات المنشورة عن مكافحة التلوث فى 64 مصنعاً فى المنطقة وتجد أن منها 56 مصنعاً تستوفى معايير المكافحة . (أ) هل تؤيد بيانات العينة إدعاء الحكومة عند مستوى معنوية 5% ؟ (ب) هل يتغير القرار إذا كان حجم العينة 124 مع بقاء نسبة المصانع التى تستوفى المعايير كما كانت من قبل ؟

(أ) هنا $H_0 : p = 0.80$ و $H_1 : p > 0.80$. وتقع منطقة رفض H_0 إلى اليمين من 1.64 انحرافاً معيارياً عند $\alpha = 5\%$. بالنسبة للعينة .

$$\bar{p} = \frac{56}{64} = 0.88 \qquad \text{and} \qquad \sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{(0.8)(0.2)}{64}} = 0.05$$

وحيث أن :

$$z = \frac{\bar{p} - p}{\sigma_{\bar{p}}} = \frac{0.88 - 0.80}{0.05} = 1.6$$

فإنها تقع داخل منطقة القبول للفرض H_0 . وهذا يعنى أنه ليس هناك سند إحصائى لادعاء الحكومة أن $p > 0.8$ عند مستوى المعنوية 5% .

(ب) لو كان حجم العينة 124 بدلا من 64 ، ولكن بقيت $\bar{p} = 0.88$ ، فإن

$$\sigma_{\bar{p}} = \frac{(0.8)(0.2)}{124} = 0.04 \qquad \text{and} \qquad z = \frac{0.88 - 0.80}{0.04} = 2$$

وتقع هذه القيمة z داخل منطقة رفض H_0 (ولا يكون هناك دليل ضد ادعاء الحكومة أن $p > 0.8$) . لاحظ أن زيادة n (مع ثبات الأشياء الأخرى على حالها) قد رفع من احتمال قبول ادعاء الحكومة .

٥ – ١٠ أوجد احتمال قبول H_0 للمسألة (٥ – ٦) إذا كانت

(أ) $\mu = \mu_0 = 80$ (ب) $\mu = 82$ (جـ) $\mu = 84$ (د) $\mu = 85$

(هـ) $\mu = 86$ (و) $\mu = 87$

(أ) إذا كانت $\mu = \mu_0 = 80$ ، $\bar{X} = 85$ ، $\sigma = 10$ و $n = 25$

$$z = \frac{\bar{X} - \mu_0}{\sigma_{\bar{X}}} = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{85 - 80}{10/\sqrt{25}} = \frac{5}{2} = 2.5$$

احتمال قبول H_0 عند $\mu = \mu_0 = 80$ هو 0.9938 (بالكشف مقابل قيمة $z = 0.5$ فى ملحق (٢) وإضافة 0.5 إلى العدد) . وعليه فإن احتمال رفض H_0 بينما فى الواقع H_0 صحيح يساوى $1 - 0.9938$ أو 0.0062

(ب) عند $\mu = 82$

$$z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} = \frac{85 - 82}{10/\sqrt{25}} = \frac{3}{2} = 1.5$$

وعليه فاحتمال قبول H_0 بينما H_0 خاطىء يساوى 0.9332 (بالكشف عن مقابل قيمة $z = 1.5$ فى ملحق (٣) وإضافة 0.5 إلى العدد) .

(جـ) عند $\mu = 84$ ، $z = (85 - 84)/2 = 1/2$ and $\beta = 0.6915$

(د) عند $\mu = 85$ ، $z = 0$ and $\beta = 0.5$

(هـ) عند $\mu = 86$ ، $z = (85 - 86)/2 = -1/2$ and $\beta = 0.5 - 0.1915 = 0.3085$

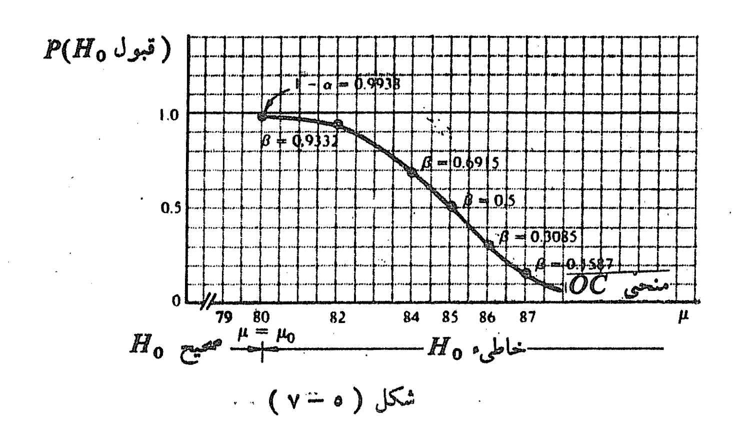
(و) عند $\mu = 87$ ، $z = -1$ and $\beta = 0.5 - 0.3413 = 0.1587$.

٥ — ١١ (أ) ارسم شكلا لإجابات المسألة (٥ — ١٠) مبيناً على المحور الرأسى احتمال قبول H_0 عندما $\mu = 80, 84, 85, 86, 87$ (ب) ماذا يوضح هذا الشكل ؟ (جـ) ما أهمية معرفة قيمة β ؟

(أ) أنظر شكل (٥ — ٧)

(ب) منحنى توصيف العمليات : OC فى شكل (٥ — ٧) يوضح قيم β عند القيم المختلفة عندما $\mu > \mu_0$. لاحظ أنه كلما زادت قيمة μ الحقيقية عن μ_0 ، كلما صغرت β (احتمال قبول H_0 عندما يكون خاطئاً) .

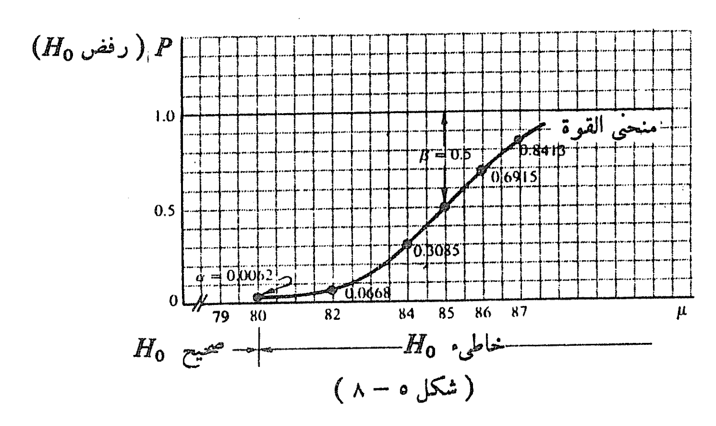
(جـ) معرفة قيمة β مهم عندما يؤدى قبول فرض خاطىء (خطأ من النوع الثانى) إلى نتائج مدمرة ، كما ، على سبيل المثال ، عند قبول عقار على أنه فعال فى حين أنه ليس كذلك . وفى مثل هذه الحالات فإننا نرغب فى أن نبقى β صغيرة ، حتى لو كان علينا قبول قيمة مرتفعة الخطأ α (خطأ من النوع الأول) . والطريقة الوحيدة لتخفيف كل من α و β معاً هو زيادة حجم العينة ، n .



شكل (٥ — ٧)

٥ — ١٢ ارسم شكلا لإجابات المسألة (٥ — ١٠) موضحاً على المحور الرأسى احتمال رفض H_0 للقيم المختلفة عندما $\mu > \mu_0$. ماذا يوضح هذا الشكل ؟ (ب) كيف كان يبدو منحنى OC فى مسألة (٥—١١) (أ) لو كان الفرض البديل $H_1 : \mu < \mu_0$

(أ) لكل قيمة $\mu > \mu_0$ ، احتمال رفض H_0 عند H_0 خاطىء يساوى β — 1 ، حيث سبق إيجاد β فى المسألة (٥—١٠) من (ب) إلى (و) . بوصل نقاط β — 1 هذه (بدءاً بقيمة α) ، نحصل على منحنى القوة (أنظر شكل ٥ — ٨) . ويوضح منحنى القوة احتمال رفض H_0 عند القيم المختلفة $\mu > \mu_0$.



(شكل ٥ – ٨)

لاحظ أنه كلما زادت μ عن μ_0 ، كلما زادت قوة الاختبار (أى ، كلما زاد احتمال رفض فرض خاطىء) .

(ب) عندما $H : \mu < \mu_0$ ، فإن منحنى OC (عند قيمة فعلية $\bar{X}$ وعند القيم البديلة المختلفة $\mu < \mu_0$) يكون مشابهاً لمنحنى القوة فى شكل (٥ – ٨) . ولكن منحنى القوة سيكون مشابهاً لمنحنى OC فى شكل (٥ – ٧) .

اختبار الفروض عن الفرق بين وسطين أو الفرق بين نسبتين :

٥ – ١٣ يرغب مشتر كبير للمصابيح الكهربائية أن يقرر ، عند مستوى معنوية 5 % ، أى صنف يشترى من بين صنفين لهما نفس السعر . لعمل هذا ، فإنه يأخذ عينة عشوائية من 100 مصباح من كل صنف فيجد أن الصنف الأول يعيش فى المتوسط $\bar{X}_1 = 980$ ساعة ، مع انحراف معيارى ، s_1 قدره 80 ساعة وبالنسبة للصنف الثانى ، $\bar{X}_2 = 1{,}010$ ساعة و $s_2 = 120$ ساعة . أى الصنفين يجب شراؤه إذا كان المشترى يرغب فى أن يصل إلى قرار عند مستوى معنوية (أ) 5% ؟ (ب) 1% ؟

(أ)

$$H_0:\ \mu_1 = \mu_2 \quad \text{or} \quad H_0:\ \mu_1 - \mu_2 = 0$$
$$H_1:\ \mu_1 \neq \mu_2 \quad \text{or} \quad H_1:\ \mu_1 - \mu_2 \neq 0$$
$$\bar{X}_1 = 980 \text{ h} \qquad s_1 = 80 \text{ h} \qquad n_1 = 100$$
$$\bar{X}_2 = 1{,}010 \text{ h} \qquad s_2 = 120 \text{ h} \qquad n_2 = 100$$

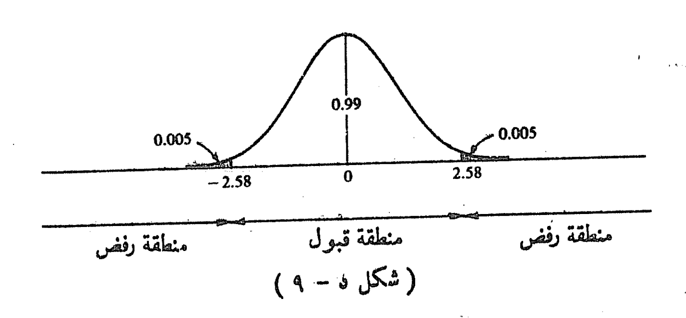
هذا اختبار ذو ذيلين وتقع منطقة القبول فى حدود ± 1.96 تحت المنحنى الطبيعى القياسى (أنظر شكل (٥ – ١) ومن ثم ،

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \cong \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{80^2}{100} + \frac{120^2}{100}} = \sqrt{64 + 144} \cong 14.42$$

$$z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{X}_1 - \bar{X}_2}} = \frac{(\bar{X}_1 - \bar{X}_2) - 0}{\sigma_{\bar{X}_1 - \bar{X}_2}} = \frac{980 - 1{,}010}{14.42} = \frac{-30}{14.42} = -2.08$$

وحيث أن قيمة z المحسوبة تقع داخل منطقة الرفض للفرض H_0 ، فعلى المشترى أن يقبل H_1 ، أى $\mu_1 \neq \mu_2$ ، عند مستوى معنوية 5% (ويفترض أنه سوف يقرر شراء الصنف الثانى) .

(ب) عند مستوى معنوية 1% فإن قيمة z المحسوبة تقع داخل منطقة القبول للفرض H_0 (انظر شكل ٥ – ٩) . ويشير هذا إلى أنه لايوجد اختلاف جوهرى بين μ_1 و μ_2 عند مستوى المعنوية 1% ، وعليه فيمكن للمشترى أن يشترى أياً من الصنفين . لاحظ أنه بالرغم أن الصنف الثانى ، يعيش أكثر من الصنف الأول إلا أن الصنف الثانى له أيضاً انحراف معيارى أكبر من الصنف الأول .



(شكل ٥ – ٩)

٥ – ١٤ متوسط الدرجات فى امتحان القبول للدراسات العليا GRE لعام ١٩٨١ لعدد 64 طالباً متقدمين الماجستير هو 640 درجة بانحراف معيارى 20 درجة . وفى عام ١٩٨٢ تقدم 81 طالباً للالتحاق بالماجستير فكان متوسط درجاتهم فى امتحان القبول 650 درجة بانحراف معيارى 40 . (أ) هل مستوى المتقدمين عام ١٩٨١ أقل من مستوى المتقدمين ١٩٨٢ عند مستوى معنوية 1% ؟ (ب) ماهى منطقة القبول بدلالة درجات امتحان GRE ؟

(أ) .

$$H_0:\ \mu_1 = \mu_2 \quad \text{and} \quad H_1:\ \mu_1 < \mu_2$$

$$\bar{X}_1 = 640 \quad s_1 = 20 \quad n_1 = 64$$

$$\bar{X}_2 = 650 \quad s_2 = 40 \quad n_2 = 81$$

وهذا اختبار الذيل الأيسر حيث تقع منطقة القبول للفرض H_0 إلى اليمين من 2.33 — تحت المنحنى الطبيعى القياسى ، وعليه .

$$\sigma_{\bar{X}_1-\bar{X}_2} \cong \sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} = \sqrt{\frac{20^2}{64}+\frac{40^2}{81}} = \sqrt{6.25+19.75} = \sqrt{26} = 5.10$$

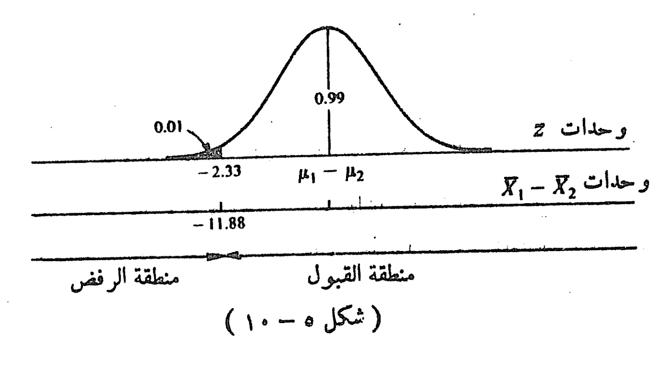
$$z = \frac{\bar{X}_1-\bar{X}_2}{\sigma_{\bar{X}_1-\bar{X}_2}} = \frac{640-650}{5.10} = \frac{-10}{5.10} = -1.96$$

وحيث أن قيمة z المحسوبة تقع داخل منطقة القبول ، تقبل H_0 . وهذا يعنى أنه لايوجد دليل إحصائى عند مستوى معنوية 1% يشير إلى أن مستوى المتقدمين يختلف بين العامين .

(ب) حيث أن الفرق المفترض بين متوسطى المجتمعين فى الفرض H_0 هو 0 ، فيمكننا إيجاد منطقة القبول للاختبار معبراً عنها بدرجات GRE كالآتى :

$$(\mu_1-\mu_2)_0 - z\sigma_{\bar{X}_1-\bar{X}_2} = 0-(2.33)(5.10) = -11.88$$

وحيث أن $\bar{X}_1 = \bar{X}_2 = 10$ ، فإنها تقع داخل منطقة القبول للفرض H_0 (أنظر شكل ٥ – ١٠) .



(شكل ٥ – ١٠)

٥ – ١٥ يرغب الاتحاد الأمريكى لطب الأسنان فى اختبار أى معجون من بين معجونى أسنان أفضل فى محاربة التسوس . أخذت عينة عشوائية من 21 شخصاً من مستعملى كل من المعجونين موضع الاختبار . ووجد أن متوسط عدد الفجوات للمجموعة الأولى على مدى 10 سنوات هو 25 بانحراف معيارى 5 وبالنسبة للمجموعة الثانية ، متوسط عدد الفجوات 23 بانحراف معيارى 5 بافتراض أن توزيع الفجوات طبيعى لمستعملى المعجون الأول والمعجون الثانى، وأن $\sigma_1^2 = \sigma_2^2$ ، حدد إذا كانت $\mu_1 = \mu_2$ عند مستوى معنوية 5%

$$H_0: \mu_1 = \mu_2 \quad \text{and} \quad H_1: \mu_1 \neq \mu_2$$

$$\bar{X}_1 = 25 \quad s_1 = 5 \quad n_1 = 21$$

$$\bar{X}_2 = 23 \quad s_2 = 4 \quad n_2 = 21$$

وحيث أن المجتمعين يتبعان التوزيع الطبيعى ولكن كلا من n_1 و n_2 أقل من 30 ومن المفترض أن $\sigma_1^2 = \sigma_2^2$ (ولكنهما غير معلومين) ، فإن توزيع المعاينة للفرق بين متوسط يتبع توزيع t بدرجات حرية $n_1 + n_2 - 2$. وحيث أنه من المفترض أن $\sigma_1^2 = \sigma_2^2$ (فيمكننا استخدام s_1^2 كتقدير σ_1^2 و s_2^2 كتقدير σ_2^2) ، فإن

$$\sigma_{\bar{X}_1 - \bar{X}_2} \cong \sqrt{\frac{s^2}{n_1} + \frac{s^2}{n_2}} \qquad (٥ - ١ أ)$$

حيث

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \qquad (٥ - ٣ أ)$$

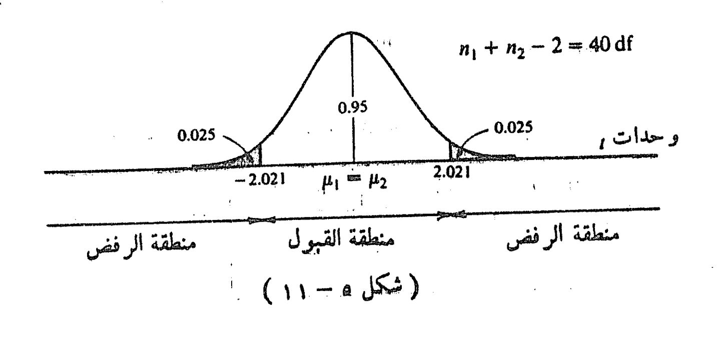
s^2 متوسط مرجح القيم s_1^2 و s_2^2 . الأوزان هى $n_1 - 1$ و $n_2 - 1$ ، كما فى معادلة (٢ – ٨ ب) المقابلة لكل من s_1^2 و s_2^2 ، للحصول على تقرير « غير متحيز » لكل من σ_1^2 و σ_2^2 (أنظر مسألة ٢ – ١٦) . وهذا اختبار ذوذيلين وتقع منطقة القبول للفرض H_0 داخل ± 2.021 تحت توزيع t مع $\alpha = 5\%$ و $n_1 + n_2 - 2 = 12 + 12 - 2 = 40$ df

$$s^2 = \frac{20(5)^2 + 20(4)^2}{40} = \frac{500 + 320}{40} = 20.5$$

$$\sigma_{\bar{X}_1 - \bar{X}_2} \cong \sqrt{\frac{20.5}{21} + \frac{20.5}{21}} = \sqrt{\frac{42}{21}} = \sqrt{2} \cong 1.41$$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sigma_{\bar{X}_1 - \bar{X}_2}} = \frac{25 - 23}{1.41} \cong 1.42$$

وحيث أن قيمة t المحسوبة تقع داخل منطقة القبول ، فإننا لانستطيع أن نرفض H_0 ، القائل بأن $\mu_1 = \mu_2$ (أنظر شكل ٥ – ١١) .



(شكل ٥ – ١١)

٥ — ١٦ افترض أن 50% من 60 مصنعاً فى إقليم ١ تخضع لمعايير مكافحة التلوث بينما 40% فقط من 40 مصنعاً فى إقليم ٢ تخضع لنفس المعايير . هل نسبة المصانع التى تخضع لمعايير مكافحة التلوث أكبر معنوياً فى إقليم ١ عنها فى إقليم ٢ عند :
(أ) مستوى المعنوية 5% ؟ (ب) مستوى المعنوية 10% ؟

(أ)

$$H_0: \; p_1 = p_2 \quad \text{and} \quad H_1: \; p_1 > p_2$$
$$\bar{p}_1 = 0.50 \quad \text{and} \quad n_1 = 60$$
$$\bar{p}_2 = 0.40 \quad \text{and} \quad n_2 = 40$$

هذا اختبار الذيل الأيمن وتقع منطقة القبول للفرض H_0 عند $\alpha = 0.05$ إلى اليسار من 1.64 تحت المنحنى الطبيعى القياسى :

$$\bar{p} = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2} = \frac{60(0.5) + 40(0.4)}{60 + 40} = \frac{30 + 16}{100} = 0.46$$

$$\sigma_{\bar{p}_1 - \bar{p}_2} = \sqrt{\frac{\bar{p}(1-\bar{p})}{n_1} + \frac{\bar{p}(1-\bar{p})}{n_2}} = \sqrt{\frac{(0.46)(0.54)}{60} + \frac{(0.46)(0.54)}{40}}$$
$$= 0.00414(0.00621) = 0.01035 = 0.10$$

وحيث أن $z = (\bar{p}_1 - \bar{p}_2)/\sigma_{\bar{p}_1 - \bar{p}_2} = (0.5 - 0.4)/0.1 = 0.10/0.10 = 1$ ، فإننا نقبل H_0 ، أى $p_1 = p_2$ عند $\alpha = 0.05$.

(ب) عند $\alpha = 0.10$ ، تقع منطقة القبول للفرض H_0 إلى اليسار من 1.28 تحت المنحنى الطبيعى القياسى . وحيث أن قيمة z المحسوبة تقع داخل منطقة القبول ، نقبل H_0 عند $\alpha = 0.10$ أيضاً .

اختبار كاى – تربيع لجودة التوفيق والاستقلال :

٥ — ١٧ أخذ مدير مصنع عينة عشوائية من 100 يوم من الأجازات المرجية ، ووجد أن 30% من القوة العاملة فى المصنع فى فئة العمر 29 — 20 قد أخذوا أجازة مرضية 26 يوماً من الإجمالى 100 يوم ، وأن 40% من القوة العاملة فى فئة العمر 39 — 30 قد أخذوا 37 يوماً ، وأن 20% فى فئة العمر 49 — 40 قد أخذوا 24 يوماً ، وأن 10% فى فئة العمر 50 فأكثر قد أخذوا 13 يوماً أجازة مرضية . كيف يمكن للمدير عند مستوى معنوية 5% أن يختبر الفرض أن العمر ليس عاملا فى أخذ أجازة مرضية ؟

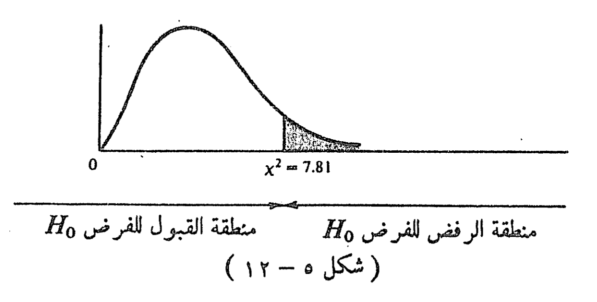
إذا كان العمر ليس عاملا ، فى أخذ أجازة مرضية ، فإن العدد المتوقع للأيام المرضية التى يأخذها العاملون فى كل فئة عمر يجب أن يكون بنفس نسبة عدد العاملين فى كل فئة عمر إلى العدد الإجمالى للعاملين بالمصنع (أنظر جدول ٥ — ٧) :

$$\chi^2 = \sum \frac{(f_0 - f_e)^2}{f_e} = \frac{(26-30)^2}{30} + \frac{(37-40)^2}{40} + \frac{(24-20)^2}{20} + \frac{(13-10)^2}{10}$$
$$= \frac{16}{30} + \frac{9}{40} + \frac{16}{20} + \frac{9}{10} \cong 2.46$$

جدول (٥ — ٧) الأجازات المرضية المشاهدة والمتوقعة

فئة العمر	20–29	30–39	40–49	50 فأكثر	الإجمالى
f_0	26	37	24	13	100
f_e	30	40	20	10	100

درجات الحرية $df = c - m - 1 = 4 - 0 - 1 = 3$. وحيث أنه لم يتم تقدير أى معلمة من المجتمع ، $m = 0$. $df = 3$ بمعنى أننا إذا عرفنا ثلاث قيم من الفئات الأربع ، فإن القيمة الرابعة ليست « حرة » أن تتغير . وحيث أن القيمة المحسوبة $\chi^2 = 2.46$ أصغر من القيمة الجدولية $\chi^2 = 7.81$ عند $\alpha = 0.05$ ودرجات حرية $df = 3$ (أنظر ملحق ٦ وشكل ٥ – ١٢) ، فإننا لانستطيع أن نرفض H_0 ، بأن العمر ليس عاملا فى أخذ أجازة مرضية . لاحظ أنه كما فى حالة توزيع t فإن هناك توزيع χ^2 مختلفاً لكل من درجات الحرية المختلفة . ولكن ، اختبار χ^2 يستخدم هنا كاختبار الذيل الأيمن فقط .



(شكل ٥ – ١٢)

٥ – ١٨ جدول (٥ – ٨) يوضح التكرارات المشاهدة والمتوقعة لأربعة أمراض نادرة (A ، B ، C ، D) فى مدينة ما . هل الفرق جوهرى بين التكرارات المتوقعة والمشاهدة للأمراض عند مستوى معنوية 10% ؟

جدول (٥ – ٨) التكرارات المشاهدة والمتوقعة لأمراض نادرة A ، B ، C و D

	نوع المرض				إجمالى
	A	B	C	D	
f_0	3	5	6	3	17
f_e	6	6	3	2	17

حيث أنه بالنسبة للأمراض C و D ، $f_e < 5$. فإننا نضم هاتين الفئتين معاً (أنظر جدول ٥ – ٩) :

$$\chi^2 = \sum \frac{(f_0 - f_e)^2}{f_e} = \frac{(3-6)^2}{6} + \frac{(5-6)^2}{6} + \frac{(9-5)^2}{5} = \frac{9}{6} + \frac{1}{6} + \frac{16}{5} = 4.87$$

وحيث أن قيمة χ^2 المحسوبة تزيد عن القيمة الجدولية $\chi^2 = 4.61$ عند $\alpha = 0.10$ و $df = 2$ فإننا نرفض H_0 ونقبل الفرض البديل H_1 ، من أن هناك فرقاً معنوياً بين التكرارات المشاهدة والمتوقعة لحدوث هذه الأمراض فى هذه المدينة. لاحظ أنه عندما $f_0 = f_e$ فإن $\chi^2 = 0$. وكلما زاد الفرق بين f_0 و f_e ، كلما كبرت قيمة χ^2 وزاد احتمال رفض H_0 . لاحظ أيضاً أنه كنتيجة لعملية التربيع فإن χ^2 لايمكن أن تكون سالبة .

جدول (٥ – ٩) التكرارات المشاهدة والمتوقعة للأمراض النادرة A ، B ، C ، و D

	نوع المرض			إجمالى
	A	B	C و D	
f_0	3	5	9	17
f_e	6	6	5	17

٥ - ١٩ جدول (٥ - ١٠) يعطى توزيع القبول لعدد 100 طالب فى 3 كليات . بمستوى معنوية 5% اختبر معنوية أن توزيع القبول هو تقريباً ذو الحدين إذا كان احتمال قبول طالب فى كلية ما 0.40

جدول (٥ - ١٠) توزيع القبول لمائة طالب فى ثلاث كليات

مرات القبـــول	عدد الطلاب
0	25
1	34
2	31
3	10
	100

احتمالات ذى الحدين الموضحة فى جدول (٥ - ١١) المناظرة لمرات قبول 0 ، 1 ، 2 أو 3 لأى طالب عند $p = 0.4$ تم الحصول عليها من ملحق ١ . وعليه

$$\chi^2 = \sum \frac{(f_0 - f_e)^2}{f_e} = \frac{(25-22)^2}{22} + \frac{(34-43)^2}{43} + \frac{(31-29)^2}{29} + \frac{(10-6)^2}{6} = \frac{9}{22} + \frac{81}{43} + \frac{2}{29} + \frac{16}{6} = 5.03$$

وحيث أن القيمة المحسوبة $\chi^2 = 5.03$ أصغر من القيمة الجدولية $\chi^2 = 7.81$ عند $\alpha = 0.05$ ودرجات حرية 3 ، فإننا لانستطيع أن نرفض H_0 ، بأن توزيع القبول يتبع توزيع ذى الحدين ، عند $p = 0.40$. لاحظ أن توزيع χ^2 هو توزيع متصل (كالتوزيع الطبيعى وتوزيع t) .

جدول (٥ - ١١) ، التكرارات المشاهدة ، احتمالات ذى الحدين ، والتكرارات المتوقعة للقبول

عدد مرات القبول	التكرارات المشاهده	احتمالات ذى الحدين		عدد المتقدمين	التكرار المتوقع للقبول
0	25	0.216	×	100	22
1	34	0.432	×	100	43
2	31	0.288	×	100	29
3	10	0.064	×	100	6
		1.000			100

٥ - ٢٠ يعطى جدول (٥ - ١٢) توزيع درجات اختبار القدرات الدراسية SAT لعينة عشوائية من 100 طالب جامعى . باستخدام مستوى معنوية 5 % اختبر ما إذا كانت درجات SAT تتبع التوزيع الطبيعى .

جدول (٥ - ١٢) التوزيع التكرارى لدرجات SAT

درجـــات SAT	عدد الطلاب
251-350	3
351-450	25
451-550	50
551-650	20
651-750	2
	100

لإجراء هذا الاختبار ، يجب أولا حساب $\bar{X}$ و s لهذا التوزيع ، كما هو موضح بجدول (٥ – ١٣) :

$$\bar{X} = \frac{\sum fX}{n} = \frac{49,300}{100} = 493$$

$$s = \sqrt{\frac{\sum fX^2 - n\bar{X}^2}{n-1}} = \sqrt{\frac{24,950,000 - (100)(493)^2}{99}} \cong 80.72$$

جدول (٥ – ١٣) حساب $\bar{X}$ و s لدرجات SAT

الفـئة	التكـرار f_0	مركز الفئة X	fX	X^2	fX^2
251–350	3	300	900	90,000	270,000
351–450	25	400	10,000	160,000	4,000,000
451–550	50	500	25,000	250,000	12,500,000
551–650	20	600	12,000	360,000	7,200,000
651–750	2	700	1,400	490,000	980,000
	100		49,300		24,950,000

فإذا كانت درجات SAT تتبع التوزيع الطبيعى ، تقدر f_e كما هو موضح فى جدول (٥ – ١٤) :

$$\chi^2 = \sum \frac{(f_0 - f_e)^2}{f_e} = \frac{(28 - 29.81)^2}{29.81} + \frac{(50 - 46.31)^2}{46.31} + \frac{(22 - 23.88)^2}{23.88} \cong 0.54$$

لاحظ أنه قد تم إدماج تكرارات أول فئتين وآخر فئتين كل فى فئة واحدة لأن $f_e < 5$. درجات الحرية $\text{df} = c - m - 1 = 5 - 2 - 1 = 2$. ولأنه تم تقدير معلمتين للمجتمع (μ و σ تم تقديرهما باستخدام $\bar{X}$ و s على الترتيب) ، فإن $m = 2$. قيمة χ^2 الجدولية عند $\alpha = 0.05$ و $\text{df} = 2$ هى 5.99 .

جدول (٥ – ١٤) التكرارات المتوقعة لدرجات SAT باستخدام $\bar{X} = 493$ و $s = 80.72$

درجات SAT الحد الأعلى للفـئة = x	$z = \frac{X - 493}{80.72}$	المساحة يسار X	مساحة المقطع	التكرار المتوقع f_e
⩽ 350	−1.77	0.0384	0.0384 × 100 = 3.84	29.81
450	−0.53	0.2981	0.2597 × 100 = 25.97	
550	0.71	0.7612	0.4631 × 100 = 46.31	46.31
650	1.94	0.9738	0.2126 × 100 = 21.26	23.88
> 750	3.18	1.0000	0.0262 × 100 = 2.62	
			1.0000 100.00	

وحيث أن قيمة χ^2 المحسوبة أصغر من القيمة الجدولية : فإنه لايمكننا رفض H_0 . أى ، لايمكننا رفض الفرض القائل بأن العينة العشوائية لدرجات SAT تأتى من توزيع طبيعى بمتوسط $\mu = 493$ و $\sigma = 80.72$.

٥ – ٢١ يوضح جدول (٥ – ١٥) للاقتران عدد النوبات القلبية التى تعرض لها الذكور والإناث فى فئات العمر المختلفة فى مدينة ما . باستخدام مستوى معنوية 1% اختبر الفرض أن العمر والجنس مستقلان فيما يتعلق بحدوث النوبات القلبية .

جدول (٥ — ١٥) عدد النوبات القلبية للذكور والإناث فى فئات العمر المختلفة فى إحدى المدن

فئة العمر	ذكور	إناث	إجمالى
أقل من 30	10	10	20
من 30 إلى 60	50	30	80
أكثر من 60	30	20	50
	90	60	150

لاختبار هذا الفرض ، يجب تقدير التكرارات المتوقعة f_e (أنظر جدول ٥ — ١٦) :

للخلية فى الصف الأول r ، والعمود الأول c

$$f_e = \frac{\sum r \sum c}{n} = \frac{(20)(90)}{150} = 12$$

للخلية فى الصف الثانى والعمود الأول

$$f_e = \frac{\sum r \sum c}{n} = \frac{(80)(90)}{150} = 48$$

ويمكن الحصول على باقى التكرارات المتوقعة بالطرح من مجموع الصف أو العمود المناظر ، وعليه

$$\chi^2 = \sum \frac{(f_0 - f_e)^2}{f_e} = \frac{(10-12)^2}{12} + \frac{(10-8)^2}{8} + \frac{(50-48)^2}{48}$$

$$+ \frac{(30-32)^2}{32} + \frac{(30-30)^2}{30} + \frac{(20-20)^2}{20} = 1.04$$

درجات الحرية $\text{df} = (r-1)(c-1) = (3-1)(2-1) = 2$ (المناظرة للتكرارات المتوقعة التى قنا بحسابها باستخدام المعادلة) . من ملحق ٦ ، $\chi^2 = 9.21$ عند $\alpha = 0.01$ ودرجات حرية 2 . وحيث أن χ^2 المحسوبة أصغر من χ^2 الجدولية ، نقبل الفرض العدمى ، H_0 ، أن العمر مستقل عن الجنس فى حدوث النوبات القلبية . ولكن هذا الاتجاه لا يختلف معنوياً مع العمر عند مستوى معنوية 1% .

جدول (٥ — ١٦) التكرارات المتوقعة للنوبات القلبية

فئة العمر	ذكور	إناث	إجمالى
أقل من 30	12	8	20
من 30 إلى 60	48	32	80
أكثر من 60	30	20	50
	90	60	150

٥ — ٢٢ أعطت عينة عشوائية من 37 عاملا فوق سن 65 فى مدينة ما النتائج الواردة فى جدول الاقتران (٥ — ١٧) . باستخدام مستوى المعنوية 10% اختبر الفرض بأن عدد الإناث والذكور من العاملين ، فى مجموعات السن 70 — 66 و 71 فأكثر ، فى المدينة مستقل عن الجنس .

جدول (٥ — ١٧) العاملون من الذكور والإناث فوق سن 65 فى مدينة

فئة العمر	إناث	ذكور	إجمالى
66 — 70	17	9	26
71 فأكثر	3	8	11
	20	17	37

جدول (٥ — ١٨) يعطى التكرارات المتوقعة . بالنسبة للخلية الأولى ،

$$f_e = \frac{\sum r \sum c}{n} = \frac{(26)(20)}{37} = 14$$

بالنسبة لباقى الخلايا ، يمكن إيجاد f_e بالطرح من مجموع الصف والعمود $1 = (1 - 2) = (1 - c)(1 - r) = \text{df}$
وحيث أن $\text{df} = 1$ و $n < 50$ ، فيجب استخدام معامل تصحيح لحساب χ^2 ، كما فى معادلة (٥ — ٤ أ) :

$$\chi^2 = \sum \frac{(|f_0 - f_e| - 0.5)^2}{f_e} \qquad (٥ - ٤ \text{ أ})$$

بالتالى

$$\chi^2 = \frac{(|17 - 14| - 0.5)^2}{14} + \frac{(|9 - 12| - 0.5)^2}{12} + \frac{(|3 - 6| - 0.5)^2}{6} + \frac{(|8 - 5| - 0.5)^2}{5}$$

$$= \frac{2.5^2}{14} + \frac{2.5^2}{12} + \frac{2.5^2}{6} + \frac{2.5^2}{5} = 3.25$$

وحيث أن قيمة χ^2 المحسوبة أكبر من قيمة الجدولية عند $\alpha = 0.10$ ودرجات حرية $\text{df} = 1$ ، فإننا نرفض الفرض H_0 ، بأن الذكور والإناث فوق سن 65 يستمرون فى العمل بصورة مستقلة عما إذا كانوا فوق أو تحت سن 70 فى هذه المدينة . أن نسبة العاملين أعلى بدرجة جوهرية للذكور فى فئة السن 70 — 66 وللإناث فى فئة السن 71 فأكثر . لاحظ أن نفس التعديل المشار إليه فى معادلة (٥ — ٤ أ) يجب إجرؤة أيضاً عند اختبار جودة التوفيق فى حالة $\text{df} = 1$ و $n < 50$.

جدول (٥ — ١٨) العدد المتوقع للعاملين من الذكور والإناث فوق سن 65

فئة العمر	إناث	ذكور	إجمالى
66–70	14	12	26
71 فأكثر	6	5	11
	20	17	37

تحليل التباين :

٥ — ٢٣ يعطى جدول (٥ — ١٩) إنتاج 8 سنوات لمزرعة تجريبية باستخدام 4 أسمدة . بافتراض أن الإنتاج باستخدام كل سماد يتبع التوزيع الطبيعى مع تساوى التباين .

جدول (٥ – ١٩) 8 سنوات باستخدام 4 أسمدة مختلفة

سماد ١	سماد ٢	سماد ٣	سماد ٤
51	47	57	50
47	50	48	61
56	58	52	57
52	61	60	65
57	51	61	58
59	48	57	53
58	59	51	61
60	50	46	59
440	424	432	464

(أ) أوجد متوسط الإنتاج لكل سماد والمتوسط الكبير لكل السنوات للأسمدة الأربعة .

(ب) قدر تباين المجتمع باستخدام التباين بين المتوسطات أو الأعمدة

(جـ) قدر تباين المجتمع من التباين داخل العينات أو الأعمدة

(د) اختبر الفرض بأن متوسطات المجتمع متساوية عند مستوى معنوية 5%

(أ)

$$\bar{X}_1 = \frac{\sum_i X_{i1}}{r} = \frac{440}{8} = 55 \qquad \bar{X}_2 = \frac{\sum_i X_{i2}}{r} = \frac{424}{8} = 53$$

$$\bar{X}_3 = \frac{\sum_i X_{i3}}{r} = \frac{432}{8} = 54 \qquad \bar{X}_4 = \frac{\sum_i X_{i4}}{r} = \frac{464}{8} = 58$$

$$\bar{\bar{X}} = \frac{\sum_J \sum_i X_{iJ}}{rc} = \frac{440 + 424 + 432 + 464}{(8)(4)} = 55$$

(ب) (من معادلات (٤ – ٢ أ) ، (٤ – ٩ أ) ، (٢ – ٧٨)

$$\sigma^2 = \frac{\sigma_{\bar{X}}^2}{n} \cong \frac{\sum (X - \bar{X})^2/(n-1)}{n}$$

وهنا

$$\sigma^2 = \frac{\sigma_{\bar{X}}^2}{n} \cong \frac{r\sum\left(\bar{X}_J - \bar{\bar{X}}\right)^2}{c-1}$$

حيث $\bar{X}_J$ هو متوسط عينة أو متوسط العمود ، $\bar{\bar{X}}$ هو المتوسط الكثير r عدد المشاهد ات فى كل عينة و c عدد العينات .

$$\sum\left(\bar{X}_J - \bar{\bar{X}}\right)^2 = (55-55)^2 + (53-55)^2 + (54-55)^2 + (58-55)^2 = 14$$

$$\sigma^2 = \frac{r\sum\left(\bar{X}_J - \bar{\bar{X}}\right)^2}{c-1} = \frac{8(14)}{3} = \frac{112}{3} = 37.33$$

وهى تقدير لتباين المجتمع من التباين بين المتوسطات أو الأعمدة .

(جـ) تقدير تباين المجتمع من التباين داخل العينات أو الأعمدة بأخذ متوسط التباينات الأربعة :

$$S_1^2 = \frac{\sum (X_{i1} - \bar{X}_1)^2}{r-1} = \frac{(51-55)^2 + (47-55)^2 + \cdots + (60-55)^2}{8-1} = \frac{144}{7} \cong 20.57$$

$$S_2^2 = \frac{\sum (X_{i2} - \bar{X}_2)^2}{r-1} = \frac{(47-53)^2 + (50-53)^2 + \cdots + (50-53)^2}{8-1} = \frac{208}{7} \cong 29.71$$

$$S_3^2 = \frac{\sum (X_{i3} - \bar{X}_3)^2}{r-1} = \frac{(57-54)^2 + (48-54)^2 + \cdots + (46-54)^2}{8-1} = \frac{216}{7} \cong 30.86$$

$$S_4^2 = \frac{\sum (X_{i4} - \bar{X}_4)^2}{r-1} = \frac{(50-58)^2 + (61-58)^2 + \cdots + (59-58)^2}{8-1} = \frac{158}{7} \cong 22.57$$

$$\sigma^2 \cong \frac{S_1^2 + S_2^2 + S_3^2 + S_4^2}{4} = \frac{20.57 + 29.71 + 30.86 + 22.57}{4} \cong 25.93$$

ويمكن التعبير عما سبق بطريقة أكثر إيجازاً كالآتى :

$$\sigma^2 = \frac{\sum S_j^2}{c} = \frac{S_1^2 + S_2^2 + \cdots + S_c^2}{c}$$

$$= \frac{\frac{\sum (X_{i1} - \bar{X}_1)^2}{r-1} + \frac{\sum (X_{i2} - \bar{X}_2)^2}{r-1} + \frac{\sum (X_{i3} - \bar{X}_3)^2}{r-1} + \frac{\sum (X_{i4} - \bar{X}_4)^2}{r-1}}{c}$$

$$= \frac{\sum \sum (X_{ij} - \bar{X}_j)^2}{(r-1)c} = \frac{144 + 208 + 216 + 158}{(7)(4)} = \frac{726}{28} = 25.93$$

(د) $F = \dfrac{\text{التباين بين متوسطات العينـــات}}{\text{التباين داخـــل العينـــات}} = \dfrac{37.33}{25.93} = 1.44$

قيمة F من ملحق (٧) عند $\alpha = 0.05$ ، درجات حرية $c - 1 = 3$ فى البسط و $\text{df } 28 = c(r-1)$ فى المقام هى 2.95 . وحيث أن القيمة المحسوبة F أصغر من القيمة الجدولية ، فإننا نقبل H_0 ، بأن متوسطات المجتمع متساوية .

٥ - ٢٤ (أ) من النتائج التى حصلنا عليها فى المسألة (٥ - ٢٣) ، أوجد قيمة كل من SSA ، SSE و SST ودرجات الحرية لكل من MSA ، MSE و SSA ، SSE ، SST . ونسبة F . (ب) من نتائج (أ) كون جدول تحليل التباين ANOVA على نمط جدول (٥ - ٤) . (جـ) قم بتحليل التباين وارسم شكلا يوضح مناطق القبول والرفض للفرض H_0 .

(أ)

$$\text{SSA} = r\sum(\bar{X}_j - \bar{\bar{X}})^2 = 112 \text{ [from Prob. 5.23(}b\text{)]}$$

$$\text{SSE} = \sum\sum(X_{ij} - \bar{X}_j)^2 = 726 \text{ [from Prob. 5.23(}c\text{)]}$$

$$\text{SST} = \sum\sum(X_{ij} - \bar{\bar{X}})^2 = (51-55)^2 + (47-55)^2 + \cdots + (59-55)^2 = 838$$

$$= \text{SSA} + \text{SSE} = 112 + 726 = 838$$

درجات الحرية الى تقابل كل منها $\text{df(SSA)} = c - 1 = 4 - 1 = 3$ ،
$\text{df (SSE)} = (r-1)c = (8-1)(4) = 28$ و $\text{df(SST)} = rc - 1 = 32 - 1 = 31$ ،
وهى أيضاً مجموع درجات الحرية التى تقابل SSA زائداً درجات الحرية التى تقابل SSE

$$MSA = \frac{SSA}{c-1} = \frac{112}{3} = 37.33$$

$$MSE = \frac{SSE}{(r-1)c} = \frac{726}{28} = 25.93$$

$$F = \frac{MSA}{MSE} = \frac{37.33}{25.93} = 1.44$$

(ب) أنظر جدول (٥ - ٢٠) .

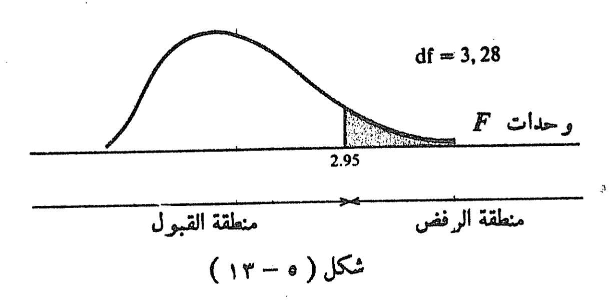
(جـ) الفروض موضع الاختبار هى

$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ مقابل $H_1 : \mu_1, \mu_2, \mu_3, \mu_4$ ليست متساوية

وحيث أن القيمة المحسوبة $F = 1.44$ أصغر من القيمة الجدولية $F = 2.95$ عند $\alpha = 0.05$ ودرجات حرية 3 و 28 ، فإننا نقبل H_0 (أنظر شكل ٥-١٣) . أى أننا نقبل الفرض العدمى H_0، بأن $\mu_1 = \mu_2 = \mu_3 = \mu_4$ وحيث أننا نعلم (من المسألة ٥ - ٢٣) أن المجتمعات تتبع التوزيع الطبيعى ولها تباينات متساوية ، فإننا ننظر إلى العينات الأربع على أنها صادرة عن نفس المجتمع . لاحظ أن MSE تقدير جيد للتباين σ^2 سواء كانت H_0 صحيحة أم لا . ولكن MSA تقريباً تساوى MSE فقط إذا كانت H_0 صحيحة (فتكون $F = 1$) . لاحظ أن توزيع F متصل وأنه يستخدم هنا لاختبار الذيل الأيمن فقط .

جدول (٥ - ٢٠) ANOVA باتجاه واحد - لتجارب الأعمدة

التغير	مجموع المربعات	درجات الحرية	متوسط المربعات	نسبة F
الذى تفسره الأعمدة (بين الأعمدة)	SSA = 112	$c - 1 = 3$	MSA = 37.33	MSA/MSE = 1.44
الخطأ أو غير المفسر (داخل الأعمدة)	SSE = 726	$(r - 1)c = 28$	MSE = 25.93	
الإجمالى	SST = 838	$rc - 1 = 31$	—	



شكل (٥ - ١٣)

٥ - ٢٥ يعطى جدول (٥ - ٢١) إنتاج مررعة تجريبية والتى استخدمت أربعة أسمدة وثلاثة مبيدات حشرية بحيث أن كل رقعة أرض كان لها فرصة متساوية فى أن تتلقى كل توليفة من نوع سماد مع نوع مبيدات حشرية (تصميم عشوائى تام) .

(أ) أوجد متوسط الإنتاج لكل سماد $\bar{X}_{.j}$ ، ولكل مبيد حشرى . X_i وللعينة ككل ، $\bar{\bar{X}}$.

(ب) أوجد إجمالى مجموع المربعات ، SST ، مجموع مربعات للأعمدة أو عامل A ، SSA ، للمبيدات الحشرية ، أو عامل B ، SSB ، وللخطأ أو البواقى غير المفسرة SSE .

(جـ) أوجد درجات الحرية لكل من SSA ، SSB ، SSE ، SST .

(د) أوجد MSA ، MSB ، MSE ، MSA/MSE ، MSB/MSE .

جدول (٥ – ٢١) الإنتاج باستخدام 4 أسمدة و 3 مبيدات حشرية

	سماد ١	سماد ٢	سماد ٣	سماد ٤
مبيد حشرى (١)	21	12	9	6
مبيد حشرى (٢)	13	10	8	5
مبيد حشرى (٣)	8	8	7	1

مبيد حشرى (٣)

(أ) متوسط العمود لكل سماد

$$\bar{X}_{.j} = \frac{\sum_i X_{ij}}{r} \qquad (٥ - ٨ أ)$$

متوسط الصف لكل مبيد حشرى

$$\bar{X}_{i.} = \frac{\sum_j X_{ij}}{c} \qquad (٥ - ٨ ب)$$

المتوسط الكبير

$$\bar{\bar{X}} = \frac{\sum \bar{X}_{i.}}{r} = \frac{\sum \bar{X}_{.j}}{c} \qquad (٥ - ٩ ب)$$

والنقاط فى رمز الدليل تشير إلى أن هناك أكثر من عامل موضع الاعتبار . النتائج موضحة فى جدول (٥ – ٢٢) .

جدول (٥ – ٢٢) الإنتاج باستخدام 4 أسمدة و 3 مبيدات حشرية (مع متوسطات الصفوف والأعمدة والمتوسطات الكبيرة)

	سماد ١	سماد ٢	سماد ٣	سماد ٤	متوسط العينة
مبيد حشرى (١)	21	12	9	6	$\bar{X}_{1.} = 12$
مبيد حشرى (٢)	13	10	8	5	$\bar{X}_{2.} = 9$
مبيد حشرى (٣)	8	8	7	1	$\bar{X}_{3.} = 6$
متوسط العينة	$\bar{X}_{.1} = 14$	$\bar{X}_{.2} = 10$	$\bar{X}_{.3} = 8$	$\bar{X}_{.4} = 4$	$\bar{\bar{X}} = 9$

(ب)

$$SST = \sum \sum (X_{ij} - \bar{\bar{X}})^2$$

$(21-9)^2 = 144$	$(12-9)^2 = 9$	$(9-9)^2 = 0$	$(6-9)^2 = 9$
$(13-9)^2 = 16$	$(10-9)^2 = 1$	$(8-9)^2 = 1$	$(5-9)^2 = 16$
$(8-9)^2 = 1$	$(8-9)^2 = 1$	$(7-9)^2 = 4$	$(1-9)^2 = 64$
161	11	5	89

$$SST = 161 + 11 + 5 + 89 = 265$$

(أ) (التغيير بين الأعمدة)

$$SSA = r\sum\left(\bar{X}_{.j} - \bar{\bar{X}}\right)^2$$

$$= 3\left[(14-9)^2 + (10-9)^2 + (8-9)^2 + (4-9)^2\right]$$

$$= 3(25 + 1 + 1 + 25) = 156$$

(ب) التغيير بين الصفوف

$$SSB = c\sum\left(\bar{X}_{i.} - \bar{\bar{X}}\right)^2$$

$$= 4\left[(12-9)^2 + (9-9)^2 + (6-9)^2\right]$$

$$= 4(9 + 0 + 9) = 72$$

$$SSE = SST - SSA - SSB = 265 - 156 - 72 = 37$$

(جـ)

$$\text{df of SSA} = c - 1 = 3$$

$$\text{df of SSB} = r - 1 = 2$$

$$\text{df of SSE} = (r-1)(c-1) = 6$$

$$\text{df of SST} = rc - 1 = 11$$

(٥ – ٥١أ)

(٥ – ١٣ب)

(٥ – ١٤أ)

(٥ – ١٥)

(د)

(٥ – ١٦)

$$MSA = \frac{SSA}{c-1} = \frac{156}{3} = 52$$

(٥ – ١٧)

$$MSB = \frac{SSB}{r-1} = \frac{72}{2} = 36$$

(٥ – ١٨)

$$MSE = \frac{SSE}{(r-1)(c-1)} = \frac{37}{11} = 3.36$$

(٥ – ١٩)

$$\frac{MSA}{MSE} = \frac{52}{3.36} = 15.48$$ نسبة F للعامل A (سماد)

(٥ – ٢٠)

$$\frac{MSB}{MSE} = \frac{36}{3.36} = 10.71$$ (نسبة F للعامل B (مبيد)

٥ – ٢٦ (أ) من نتائج مسألة (٥ – ٢٥) كوّن جدول ANOVA على نمط جدول (٥ – ٤) .

(ب) عند مستوى معنوية 1% اختبر الفرض بأن متوسطات المجتمعات للعامل A (الأعمدة) متطابقة .

(جـ) عند مستوى معنوية 1% اختبر الفرض بأن متوسطات المجتمعات للعامل B (المبيدات الحشرية) متطابقة .

(أ) أنظر جدول (٥ – ٢٣)

جدول (٥ — ٢٣) جدول ANOVA لعاملين لقياس تأثير الأسمدة والمبيدات على الإنتاج

F	متوسط المربعات	درجات الحرية	مجموع المربعات	التفسير
$\frac{MSA}{MSE} = 15.48$	$MSA = 52$	$c - 1 = 3$	$SSA = 156$	الذي تفسره الأسمدة (بين الأعمدة)
$\frac{MSB}{MSE} = 10.71$	$MSB = 36$	$r - 1 = 2$	$SSB = 72$	الذى تفسره المبيدات (بين الصفوف)
	$MSE = 3.36$	$(r-1)(c-1) = 6$	$SSE = 37$	الخطأ أو غير المفسر
	—	$rc - 1 = 11$	$SST = 265$	الإجمالى

(ب) الفروض موضع الاختبار هى :

$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ مقابل $H_1 : \mu_1, \mu_2, \mu_3, \mu_4$ ليست متساوية

حيث μ تشير إلى المتوسطات المختلفة لمجتمعات العامل A (السماد) . بالنسبة للعامل A ، $F = 9.78$ (من ملحق ٧) لدرجات حرية 3 (فى البسط) و 6 (فى المقام) و $\alpha = 0.01$. وحيث أن القيمة المحسوبة $F = 15.48$ (من جدول ٥ — ٢٣) تزيد عن القيمة الجدولية F ، فإننا نرفض H_0 ونقبل H_1 ، بأن متوسطات المجتمعات العامل A (السماد) ليست متساوية .

(جـ) المجموعة الثانية من الفروض موضع الاختبار هى

$H_0 : \mu_1 = \mu_2 = \mu_3$ مقابل $H_1 : \mu_1, \mu_2, \mu_3$ ليست متساوية حيث تشير μ هنا إلى المتوسطات المختلفة للعامل B (المبيدات) . بالنسبة للعامل B ، $F = 10.92$ (من ملحق ٧) لدرجات حرية 2 و 6 و $\alpha = 1.01$. وحيث أن القيمة المحسوبة $F = 10.71$ أصغر من القيمة الجدولية F ، فإننا نقبل H_0 ، بأن متوسطات المجتمعات للعامل B (المبيد) متساوية . لاحظ أنه فى تحليل التباين فى اتجاهين أو لعاملين (مع جدول ANOVA كجدول ٥ — ٢٣) فإنه يمكننا اختبار فرضين عدميين ، واحد لعامل A وواحد لعامل B .

٥ — ٢٧ يعطى جدول ٥ — ٢٤ دخل السنة الأولى (بآلاف الدولارات) للطلاب الحاصلين على درجة الماجستير من 5 مدارس حسب ترتيبهم عند التخرج فى 3 مجموعات . اختبر عند مستوى معنوية 5% أن المتوسطات متطابقة (أ) لمجتمعات المدارس و(ب) لمجتمعات الترتيب عند التخرج .

جدول (٥ — ٢٤) دخل السنة الأولى لخريجى الماجستير من 5 مدارس و 3 مجموعات حسب ترتيب التخرج (بآلاف الدولارات)

متوسط العينة	مدرسة ٥	مدرسة ٤	مدرسة ٣	مدرسة ٢	مدرسة ١	الترتيب فى الدفعة
$\bar{X}_{1.} = 16$	12	14	16	18	20	١/٣ الدفعة الأعلى
$\bar{X}_{2.} = 14$	10	12	13	16	19	١/٣ الدفعة الوسطى
$\bar{X}_{3.} = 12$	8	10	10	14	18	١/٣ الدفعة الدنيا
$\bar{\bar{X}} = 14$	$\bar{X}_{.5} = 10$	$\bar{X}_{.4} = 12$	$\bar{X}_{.3} = 13$	$\bar{X}_{.2} = 16$	$\bar{X}_{.1} = 19$	متوسط العينة

(أ) الفروض موضع الاختبار هي :

$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$ مقابل $H_1 : \mu_1, \mu_2, \mu_3, \mu_4, \mu_5$ ليست متساوية

حيث μ تشير إلى المتوسطات المختلفة للمجتمعات للعامل A (المدرسة) .

$$SST = \sum\sum\left(X_{ij} - \bar{\bar{X}}\right)^2$$

$(20-14)^2 = 36$	$(18-14)^2 = 16$	$(16-14)^2 = 4$	$(14-14)^2 = 0$	$(12-14)^2 = 4$
$(19-14)^2 = 25$	$(16-14)^2 = 4$	$(13-14)^2 = 1$	$(12-14)^2 = 4$	$(10-14)^2 = 16$
$(18-14)^2 = 16$	$(14-14)^2 = 0$	$(10-14)^2 = 16$	$(10-14)^2 = 16$	$(8-14)^2 = 36$
77	20	21	20	56

$$SST = 77 + 20 + 21 + 20 + 56 = 194$$

$$SSA = r\sum\left(\bar{X}_{.j} - \bar{\bar{X}}\right)^2$$ (التفسير بين الأعمدة)

$$= 3\left[(19-14)^2 + (16-14)^2 + (13-14)^2 + (12-14)^2 + (10-14)^2\right] = 3(25 + 4 + 1 + 4 + 16)$$

$$= 150$$

$$SSB = c\sum\left(\bar{X}_{i.} - \bar{\bar{X}}\right)^2 = 5\left[(16-14)^2 + (14-14)^2 + (12-14)^2\right] = 5(4 + 0 + 4) = 40$$

$$SSE = SST - SSA - SSB = 194 - 150 - 40 = 4$$

ويظهر تلخيص هذه النتائج فى جدول (٥ — ٢٥) . من ملحق ٧ $F = 3.84$ لدرجات حرية 4 و 8 و $\alpha = 0.05$. وحيث أن القيمة المحسوبة $F = 70$ فإننا نرفض H_0 ونقبل H_1 ، بأن متوسطات المجتمع لدخل السنة الأولى للمدارس الخمس مختلفة .

جدول (٥ — ٢٥) جدول ANOVA باتجاهين لدخل السنة الأولى

التفسير	مجموع المربعات	درجات الحرية	متوسط المربعات	F
الذى تفسره المدارس (A) (بين الأعمدة)	SSA = 150	$c - 1 = 4$	$MSA = \frac{150}{4} = 37.5$	$\frac{MSA}{MSE} = \frac{37.5}{0.5} = 70$
الذى يفسره الترتيب (B)(بين الصفوف)	SSB = 40	$r - 1 = 2$	$MSB = \frac{40}{2} = 20$	$\frac{MSB}{MSE} = \frac{20}{0.5} = 40$
الخطأ أو غير المفسر	SSE = 4	$(r-1)(c-1) = 8$	$MSE = \frac{4}{8} = 0.5$	
الإجمالى	SST = 194	$rc - 1 = 14$	—	

(ب) الفروض موضع الاختبار هي :

$H_0 : \mu_1 = \mu_2 = \mu_3$ مقابل $H_1 : \mu_1, \mu_2, \mu_3$ ليست متساوية

حيث تشير μ إلى متوسطات المجتمعات المختلفة للعامل B (الترتيب فى الدفعة) . من جدول (٥ — ٢٥) نحصل على القيمة المحسوبة $F = MSB / MSE = 40$. وحيث أن هذه القيمة أكبر من القيمة الجدولية $F = 4.46$

لدرجات حرية 2 و 8 و $\alpha = 0.05$ ، فإننا نرفض H_0 ونقبل H_1 ، بأن متوسطات المجتمع لدخل السنة الأولى للمجموعات الثلاث للترتيب فى الدفعة مختلفة . وعليه فإن كلا من نوع المدرسة والترتيب فى الدفعة ذا دلالة إحصائية عند مستوى معنوية 5% فى تفسير الاختلافات فى دخل السنة الأولى . ويفترض التحليل السابق ضمنياً أن تأثير العاملين قابل للإضافة (أى أنه ليس هناك تفاعل بينهما) .

مسائل اضافية

اختبار الفروض

٥ – ٢٨ (أ) ماذا تسمى خطأ قبول فرض خاطىء ؟ رفض فرض صحيح ؟

(ب) ماهو الرمز المستخدم عادة لاحتمال خطأ من النوع الأول ؟ ماهو الإسم البديل له ؟

(جـ) ماهو الرمز المستخدم عادة لاحتمال خطأ من النوع الثانى)؟

(د) ماهو مستوى الثقة ؟ (هـ) إذا خفضت α من 5 إلى 1 % ، ماذا يحدث للمعامل β ؟

الإجابة (أ) خطأ من النوع الثانى ، خطأ من النوع الأول (ب) α ، مستوى المعنوية (جـ) β (د) $1-\alpha$ (هـ (تزيد β)

٥ – ٢٩ عند مستوى معنوية 5% ، متى يكون من الأرجح أن تقبل كلية الدراسات العليا الفرض بأن متوسط درجات امتحان القبول GRE للمتقدمين : (أ) يساوى 600 ؟ (ب) أكبر من 600 ؟ (جـ) أصغر من 600 ؟

الإجابة (أ) كلما قرب متوسط العينة ، $\bar{X}$ ، من 600 (ب) كلما كانت $\bar{X} > 600$ (جـ) كلما كانت $\bar{X} < 600$

اختبار فروض عن الوسط والنسبة فى المجتمع :

٥ – ٣٠ يحتاج صاحب مصنع طائرات أن يشترى صحائف ألمنيوم بسمك 0.05 in . الصحائف الأقل سمكاً غير ملائمة والأكثر سمكاً أثقل من اللازم . يأخذ المنتج عينة عشوائية من 100 صحيفة من مورد الصحائف الألمنيوم ويجد أن متوسط سمكها 0.048 in وانحراف معيارى 0.01 in . هل يجب على المنتج شراء صحائف الألمنيوم من هذا المورد إذا كان يرغب فى أن يتخذ قراراً عند مستوى معنوية 5% ؟

الإجابة : لا

٥ – ٣١ حدد منطقة القبول لمسألة (٥ – ٣٠) بالبوصة .

الإجابة : من 0.04804 إلى 0.05196 بوصة .

٥ – ٣٢ يعرف مركز تجنيد بالبحرية من الخبرة الماضية أن أطوال المجندين موزعة طبيعياً بمتوسط ، μ ، 180 cm (1 cm=1/100 m) وانحراف معيارى μ ، 10 cm . ويرغب مركز التجنيد أن يختبر عند مستوى معنوية 1 % الفرض أن متوسط طول المجندين هذا العام أكبر من 180 cm . لعمل هذا ، يأخذ ضابط التجنيد عينة عشوائية من 46 مجنداً ويجد أن متوسط الطول فى العينة 182 cm . (أ) هل يقبل ضابط التجنيد الفرض ؟ (ب) ماهى منطقة الرفض للاختبار بالسنتيمتر ؟

الإجابة : (أ) لا (ب) أقل من 182.9125 cm .

٥ – ٣٣ ترغب مشترية لأجزاء الكترونية أن تختبر الفرض أنها ، أى الأجزاء ، تعيش أقل من 100 ساعة . . لعمل هذا فإنها تأخذ عينة عشوائية مكونة من 16 من هذه المكونات وتجد أنها فى المتوسط تعيش 96 ساعة مع انحراف معيارى 8 ساعة . فإذا كانت

المشترية تعلم أن عمر الأجزاء موزع طبقاً للتوزيع الطبيعى ، فهل يجب أن تقبل الفرض أنها تعيش لمدة أقل من 100 ساعة عند (أ) مستوى الثقة 95% ؟ (ب) مستوى الثقة 99% ؟

الإجابة : (أ) نعم (ب) لا

٥ – ٣٤ فى الماضى ، حصل 20% من المتقدمين للقبول ببرنامج ماجستير على درجات فى امتحان القبول GRE فوق 650 . من بين 88 طالباً قبلوا بالبرنامج عام ١٩٨١ ، حصل 22 طالباً على درجات GRE أعلى من 650 . هل حصل المتقدمون للبرنامج عام ١٩٨١ على درجات GRE أعلى من السنوات السابقة عند مستوى معنوية 5% ؟

الإجابة : لا

٥ – ٣٥ أوجد احتمال قبول H_0 (أن $\mu = 650$) للمسألة ٥ – ٣٤ إذا كانت $\sigma_p = 0.043$ و (أ) $p = 0.20$ (ب) $p = 0.22$ (جـ) $p = 0.24$ (د) $p = 0.25$ (هـ) $p = 9.26$ (و) $p = 0.28$.

الإجابة : (أ) 0.877 (ب) 0.758 (جـ) 0.591 (د) 0.5 (هـ) 0.409 (و) 0.242

٥ – ٣٦ ماقيمة α عند $p = 0.20$ فى المسألة (٥ – ٣٤) (ب) كيف يمكن اشتقاق منحنى OC للمسألة (٥ – ٣٤) ؟

الإجابة : (أ) 0.123 (ب) بضم قيمة $1 - \alpha$ عندما $p = 0.20$ إلى قيم β السابق إيجادها فى مسألة (٥ – ٣٥) (ب) إلى (و) المقابلة للقيم المختلفة للنسبة $p > 0.20$

٥ – ٣٧ أوجد احتمال رفض H_0 (أن $\mu = 650$) للمسألة (٥ – ٣٤) إذا كانت $\sigma_p = 0.043$ و(أ) $p = 0.20$ ، (ب) $p = 0.22$ ، (جـ) $p = 0.24$ ، (د) $p = 0.25$ (هـ) $p = 0.26$ ، (و) $p = 0.28$

الإجابة : (أ) 0.123 (ب) 0.242 (جـ) 0.409 (د) 0.5 (هـ) 0.591 (و) 0.758

٥ – ٣٨ كيف يمكن الحصول على منحنى القوة للمسألة (٥ – ٣٤) ؟

الإجابة : بضم القيم السابق الحصول عليها فى المسألة (٥ – ٣٧) من (أ) إلى (و) المناظرة للقيم المختلفة للنسبة $p > 0.2$

اختبار الفروع عن الفرق بين وسطين أو الفرق بين نسبتين :

٥ – ٣٩ ترغب شركة استثمارات أن تقرر بمستوى معنوية 5% إذا كانت أجور عمال البناء تختلف جوهرياً فى نيويورك عنها فى شيكاغو . وقد أعطت عينة عشوائية من 100 عامل بناء فى نيويورك متوسط أجر أسبوعى قدره 400$ مع انحراف معيارى قدره 100$ وفى شيكاغو ، أعطت عينة عشوائية من 75 عامل متوسط أجر أسبوعى قدره 375$ مع انحراف معيارى قدره 80$. هل هناك فرق معنوى بين أجور عمال البناء فى نيويورك وشيكاغو عند (أ) مستوى 5% ؟ (ب) مستوى 10% ؟

الإجابة : (أ) لا (ب) نعم

٥ – ٤٠ عينة عشوائية من 21 لاعباً من فريق AFC للرجبى أعطت متوسطاً لوزن اللاعبين قدره 265 lb مع انحراف معيارى قدره 30 lb بينما عينة عشوائية من 11 لاعباً من فريق NFC أعطت متوسط وزن قدره 240 lb مع انحراف معيارى قدره 20 lb . هل متوسط الوزن لكل لاعبى فريق AFC أعلى من المتوسط لفريق NFC عند مستوى معنوية 1% ؟

الإجابة : نعم

٥ – ٤١ عينة عشوائية من 100 جندى تشير إلى أن 20% تزوجوا فى السنة ١ . بينما 30% تزوجوا فى السنة . حدد هل يقبل الفرض أن نسبة الذين تزوجوا فى السنة ١ أصغر من الذين تزوجوا فى السنة ٢ (أ) عند مستوى معنوية 5% (ب) عند مستوى معنوية 1% .

الإجابة : (أ) اقبل الفرض (ب) ارفض الفرض .

اختبارات كاى – تربيع لجودة التوفيق والاستقلال :

٥ – ٤٢ ألقيت نردة 60 مرة فأعطت النتائج التالية : ظهر العدد 1 ، 12 مرة ، ظهر العدد 2 ، 8 مرات ، ظهر العدد 3 ، 13 مرة ، ظهر العدد 4 ، 12 مرة ، ظهر العدد 5 ، 7 مرات ، وظهر العدد 6 ، 8 مرات . هل النردة متوازنة عند مستوى معنوية 5%

الإجابة : نعم

٥ – ٤٣ يحتوى وعاء على كرات من أربعة ألوان : الأخضر ، الأبيض ، الأحمر ، والأزرق . التقطت كرة من الوعاء وسجل لونها . ثم أعيدت الكرة إلى الوعاء ، وخلطت الكرات جيداً والتقطت كرة أخرى . وأعيدت هذه العملية 18 مرة وكانت النتيجة أن ظهرت كرة خضراء 8 مرات ، وظهرت كرة بيضاء 7 مرات ، وظهرت كرة حمراء مرة واحدة ، وظهرت كرة زرقاء مرتين . هل يحتوى الوعاء على أعداد متساوية من الكرات الخضراء والبيضاء والحمراء أو الزرقاء ؟ اختبر الفرض عند مستوى معنوية 5% .

الإجابة : يجب قبول الفرض عند مستوى معنوية 5% أن الوعاء يحتوى على كرات منها 6 خضراء ، 6 بيضاء ، 6 حمراء أو زرقاء .

٥ – ٤٤ تشير عينة عشوائية من 64 مدينة فى الولايات المتحدة أن عدد الأيام الممطرة أثناء شهر يونيو كما فى جدول (٥ – ٢٦) هل تتبع الأيام الممطرة فى مدن الولايات المتحدة التوزيع الطبيعى بوسط $\mu = 3$ وانحراف معيارى $\sigma = 2$ عند مستوى المعنوية 10% ؟

الإجابة : لا .

جدول (٥ – ٢٦) عدد الأيام الممطرة خلال شهر يونيو
فى 64 مدينة أمريكية

عدد المدن	عدد الأيام الممطرة
10	0
12	1
22	2
13	3
6	4
1	5
64	

٥ – ٤٥ يعطى جدول الاقتران ٥ – ٢٧ عدد الأجزاء الألكترونية المقبولة وغير المقبولة المنتجة خلال ساعات الصباح المختلفة فى عينة عشوائية من إنتاج المصنع . هل يجب قبول أو رفض الفرض عند مستوى معنوية 5% بأن إنتاج الوحدات المقبولة مستقل عن ساعة الصباح التى أنتج خلالها ؟

الإجابة : أقبل H_0

جدول (٥–٢٧) الوحدات المقبولة وغير المقبولة من الأجزاء المنتجة خلال ساعات الصباح

	الإجمالى	١١–١٢ صباحاً	١٠–١١ صباحاً	٩ – ١٠ صباحاً	٨ – ٩ صباحاً
مقبولة	280	65	80	75	60
غير مقبولة	120	35	30	25	30
	400	100	110	100	90

٥ – ٤٦ يعطى جدول الاقتران (٥–٢٨) عدد الناخبين الذين صوتوا لصالح الديمقراطيين أو الجمهوريين تحت سن 40 وسن 40 فأكثر فى عينة عشوائية من 30 ناخباً فى إحدى المدن . هل التصويت للديمقراطيين أو الجمهوريين مستقل عما إذا كان الناخب تحت سن 40 ، أو فوقها فى هذه المدينة عند مستوى معنوية 5% ؟

الإجابة : لا

جدول (٥ – ٢٨) الديمقراطيين والجمهوريين عند سن أقل من 40 وسن 40 فأكثر

إجمالى	الجمهوريين	الديمقراطيين	فئة السن
11	5	6	تحت سن ٤٠
19	9	10	٤٠ فأكثر
30	14	16	

تحليل التباين :

٥ – ٤٧ يعطى جدول (٥–٢٩) عدد الأميال فى الجالون لأربعة أنواع من الأوكتين فى البنزين لمدة 5 أيام . افترض أن عدد الأميال للجالون لكل نوع أوكتين موزع طبيعياً مع تساوى التباين . هل يجب قبول أم رفض الفرض بأن متوسطات المجتمع متساوية عند مستوى معنوية 5 % ؟

الإجابة : رفض

جدول (٥ – ٢٩) عدد الأميال للجالون لأربعة أنواع من البنزين لمدة 5 أيام

النوع ٤	النوع ٣	النوع ٢	النوع ١
17	16	12	12
15	14	14	11
17	15	13	12
16	13	15	13
18	14	14	11

٥ – ٤٨ يعطى جدول (٥ – ٣٠) عدد الأميال للجالون لأربعة أنواع من الأوكتين فى البنزين وثلاثة أنواع من السيارات (ثقيلة ومتوسطة وخفيفة) فى تصميم عشوائى تام . هل يجب عند مستوى معنوية 1% قبول أو رفض الفرض أن متوسطات المجتمع متساوية لكل :
(أ) أنواع الأوكتين فى البنزين ؟ (ب) أنواع السيارات ؟

الإجابة : (أ) نعم (ب) لا .

جدول (٥ – ٣٠) عدد الأميال للجالون الواحد لأربعة أنواع أوكتين وثلاثة أنواع سيارات

أوكتين ٤	أوكتين ٣	أوكتين ٢	أوكتين ١	نوع السيارة
10	9	9	8	ثقيلة
17	18	15	16	متوسطة
30	28	26	24	خفيفة

٥ – ٤٩ يعطى جدول (٥ – ٣١) بيانات مبيعات الصابون لثلاثة أغلفة مختلفة وأربعة تركيبات مختلفة فى تصميم عشوائى تام . هل يجب عند مستوى معنوية 5 % قبول أو رفض الفرض أن متوسطات المجتمع متساوية لكل : (أ) غلاف ؟ (ب) تركيبة ؟

الإجابة : (أ) لا (ب) نعم .

جدول (٥ – ٣١) مبيعات الصابون لثلاثة أغلفة وأربعة تركيبات

	غلاف ١	غلاف ٢	غلاف ٣
تركيبة (١)	87	78	90
تركيبة (٢)	79	79	84
تركيبة (٣)	83	81	91
تركيبة (٤)	85	83	89

امتحان احصاء

١ - يعطى جدول ١ التوزيع التكرارى لمعدلات البطالة فى عينة من 20 مدينة كبيرة فى الولايات المتحدة عام ١٩٨٠ .

(أ) أوجد الوسط الحسابى والوسيط والمنوال لمعدلات البطالة .

(ب) أوجد التباين والانحراف المعيارى ومعامل الاختلاف .

(جـ) أوجد معامل بيرسون للالتواء وارسم المدرج التكرارى النسبى .

جدول (١) التوزيع التكرارى لمعدلات البطالة

التكـرار	معـدل البطـالة
2	7.0–7.4
4	7.5–7.9
5	8.0–8.4
4	8.5–8.9
3	9.0–9.4
2	9.5–9.9
$n = 20$	

٢ - من المعروف أن عمر أحد الأجزاء الألكترونية يتبع التوزيع الطبيعى بمتوسط 1,000 ساعة وانحراف معيارى 80 ساعة . ما احتمال أن يكون عمر جزء مسحوب عشوائياً من خط الإنتاج (أ) بين 1,120 و 1,180 ساعة ؟ (ب) بين 955 و 975 ساعة ؟ (جـ) أقل من 955 ساعة ؟ (د) فوق 975 ساعة ؟ (هـ) ارسم التوزيع الطبيعى والتوزيع الطبيعى القياسى لهذه المسألة وظلل المساحة المناظرة للجزء (د) .

٣ - متوسط درجات اختبار IQ (اختبار الذكاء) لعينة عشوائية من 25 طالباً فى جامعة ماهو 110 . فإذا كان المعروف أن توزيع درجات IQ فى الجامعة يتبع التوزيع الطبيعى بانحراف معيارى 10 (أ) أوجد فترة الثقة 95% للوسط غير المعلوم لدرجات مجتمع الطلاب فى الجامعة (ب) أجب عن نفس السؤال إذا كان الانحراف المعيارى للمجتمع غير معلوم ، ولكن بحساب الانحراف المعيارى من العينة وجد أنه يساوى 8 (جـ) حدد كل الحالات الممكنة التى يمكن عندها استخدام التوزيع الطبيعى ، توزيع t ، أو نظرية تشبتشيف .

٤ - تبيع شركة مسحوق صابون معبأ فى مصنعين . وتعلم الشركة من الخبرة الماضية أن كمية المسحوق فى الصناديق المعبأة فى المصنعين تتبع التوزيع الطبيعى . أخذت الشركة عينة عشوائية من 25 صندوقاً من إنتاج كل مصنع فوجدت أن المتوسط والانحراف المعيارى للوزن فى الصناديق من المصنع ١ هو 1,064 جرام (2.34 lb) و 100 جرام على الترتيب . وبالنسبة للعينة من مصنع ٢ كان المتوسط 1,024 جرام والانحراف المعيارى 60 جرام (أ) هل يمكن أن تدعى بدرجة ثقة 95% أن صناديق الصابون من مصنع ١ تحتوى على أكثر من 1,000 جرام ؟ (ب) اختبر عند مستوى ثقة 95% أن كمية الصابون فى الصناديق من المصنعين متساوية .

الإجابة :

١ - (أ) أنظر جدول ٢ .

$$\bar{X} = \frac{\sum fX}{n} = \frac{168.0}{20} = 8.4\%$$

$$\text{Med} = L + \frac{n/2 - F}{f_m} c = 8.0 + \frac{20/2 - 6}{5} 0.4 = 8.32\%$$

$$\text{Mode} = L + \frac{d_1}{d_1 + d_2} c = 8.0 + \frac{1}{1+1} 0.4 = 8.2\%$$

جدول (٢) حسابات إيجاد الوسط الحسابي والوسيط والمنوال

معدل البطالة ، نسبة مئـــوية	مركز الفئـــة X	التكـــرار f	fX
7.0–7.4	7.2	2	14.4
7.5–7.9	7.7	4	30.8
8.0–8.4	8.2	5	41.0
8.5–8.9	8.7	4	34.8
9.0–9.4	9.2	3	27.6
9.5–9.9	9.7	2	19.4
		$\sum f = n = 20$	$\sum fX = 168.0$

(ب) أنظر جدول (٣) .

$$s^2 = \frac{\sum f(X - \bar{X})^2}{n-1} = \frac{10.70}{19} \cong 0.56\% \text{ مربعـــة}$$

$$s = \sqrt{s^2} \cong 0.75\%$$

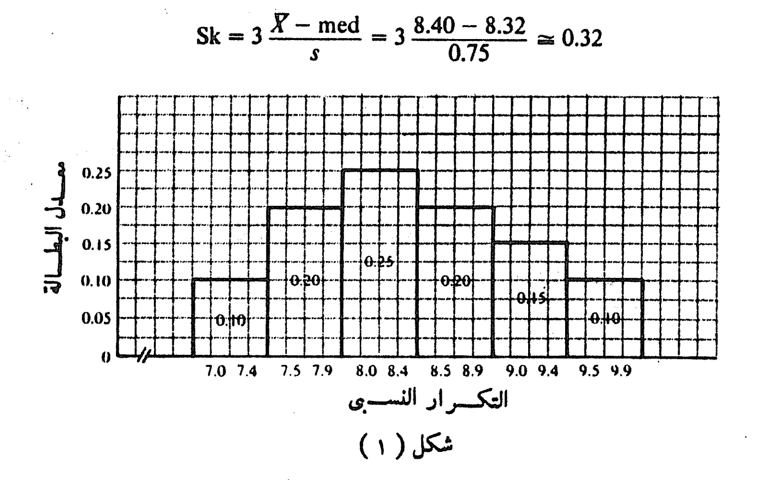
$$V = \frac{s}{\bar{X}} = \frac{0.75\%}{8.4\%} \cong 0.09$$

جدول (٣) حسابات إيجاد التباين والانحراف المعيارى ومعامل الاختلاف

معدل البطالة	مركز الفئة X	التكرار f	متوسط $\bar{X}$	$(X - \bar{X})$	$(X - \bar{X})^2$	$f(X - \bar{X})^2$
7.0–7.4	7.2	2	8.4	− 1.2	1.44	2.88
7.5–7.9	7.7	4	8.4	− 0.7	0.49	1.96
8.0–8.4	8.2	5	8.4	− 0.2	0.04	0.20
8.5–8.9	8.7	4	8.4	0.3	0.09	0.36
9.0–9.4	9.2	3	8.4	0.8	0.64	1.92
9.5–9.9	9.7	2	8.4	1.3	1.69	3.38
		$\sum f = n = 20$				$\sum f(X - \bar{X})^2 = 10.70$

(جـ) (أنظر شكل ۱)

$$\text{Sk} = 3\frac{\bar{X} - \text{med}}{s} = 3\frac{8.40 - 8.32}{0.75} \cong 0.32$$



شكل (۱)

٢ - (أ) تطلب المسألة إيجاد $P(1{,}120 < X < 1{,}180)$ حيث تشير X إلى الزمن مقيساً بالساعة لعمر الأجزاء الألكترونية . بمعلومية أن $\mu = 1000$ ساعة و $\sigma = 80$ ساعة وبوضع $X_1 = 1{,}120$ و $X_2 = 1{,}180$ نحصل على

$$z_1 = \frac{X_1 - \mu}{\sigma} = \frac{1{,}120 - 1{,}000}{80} = 1.5 \quad \text{و} \quad z_2 = \frac{1{,}180 - 1{,}000}{80} = 2.25$$

وبالكشف بالجدول (جدول التوزيع الطبيعي القياسي) مقابل $z = 2.25$ نحصل على 4878 ومقابل القيمة $z = 1.5$ نحصل على 0.4332 ثم بطرح العددين نحصل على

$$P(1{,}120 < X < 1{,}180) = 0.0546, \text{ or } 5.46\%$$

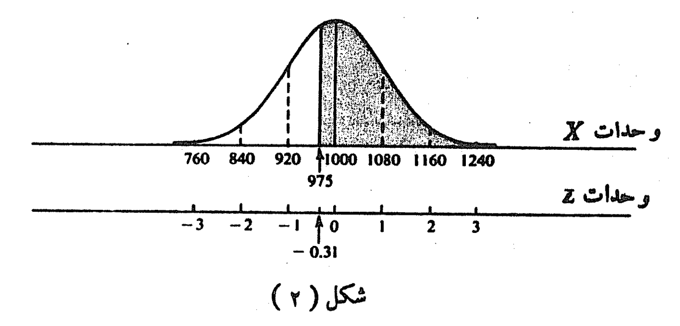
(ب)
$$z_1 = \frac{955 - 1{,}000}{80} = -0.5625 \quad \text{و} \quad z_2 = \frac{975 - 1{,}000}{80} = -0.3125$$

بالكشف بالجدول مقابل $z_1 = 0.56$ نحصل على 0.2123 ، ومقابل قيمة $z_2 = 0.31$ نحصل على 0.1217 . وعليه تكون $P(955 < X < 975) = 0.2123 - 0.1217 = 0.0906$ أو 9.06%

(جـ) $P(X < 955) = 0.5 - 0.2123 = 0.2877, \text{ or } 28.77\%.$

(د) $P(X > 975) = 0.1217 + 0.5 = 0.6217, \text{ or } 62.17\%.$

(هـ) أنظر شكل (٢) .



شكل (٢)

٣ - (أ) حيث أن المجتمع يتبع التوزيع الطبيعي و σ معلومة فيمكن استخدام التوزيع الطبيعي :

$$\mu = \bar{X} \pm z\sigma_{\bar{X}} = \bar{X} \pm z\frac{\sigma}{\sqrt{n}} = 110 \pm 1.96\frac{10}{\sqrt{25}} = 110 \pm 3.92$$

فتكون μ بين 106.08 و 113.92 بدرجة ثقة 95% .

(ب) حيث أن التوزيع طبيعي ، $n < 30$ و σ غير معلومة ، فيجب استخدام توزيع t بدلا من التوزيع الطبيعي ، مع استخدام s كتقدير بدلا من σ

$$\mu = \bar{X} \pm t_{0.025}\frac{s}{\sqrt{n}} \qquad t_{0.025} \text{ with 24 df} = 2.064$$

$$= 110 \pm 2.064\frac{8}{\sqrt{25}}$$

$$= 110 \pm 3.30$$

وتكون μ بين 106.70 و113.30 بدرجة ثقة 95% .

(ج) يمكن استخدام التوزيع الطبيعي (١) إذا كان المجتمع الأصلي طبيعياً ، σ و $n < 30$ معلومة (٢) إذا كانت $n \geq 30$ (بالجوء لنظرية النهاية المركزية) وباستخدام s كتقدير بدلا من t يمكن استخدام توزيع t (لدرجات الحرية المعينة) عندما $n < 30$ ولكن σ غير معلومة والمجتمع الذى أخذت منه العينة من المعلوم أنه يتبع التوزيع الطبيعي . وفى غير الحالات السابقة نلجأ إلى استخدام متباينة تشبتشيف أو زيادة حجم العينة إلى $n \geq 30$ (حتى يمكن استخدام التوزيع الطبيعي) .

٤ – (أ) حيث أن الشركة ترغب فى اختبار ما إذا كان $\mu > 1,000\,g$ فى مصنع ١ ، فإننا بصدد اختبار الذيل الأيمن :

$$H_0: \ \mu_1 = 1,000 \quad و \quad H_1: \ \mu_1 > 1,000$$

وحيث أن المجتمع يتبع التوزيع الطبيعي ولكن $n < 30$ و σ غير معلومة ، فيجب استخدام توزيع t بدرجات حرية $n - 1 = 24$:

$$t = \frac{\bar{X}_1 - \mu_1}{s_1/\sqrt{n_1}} = \frac{1,064 - 1,000}{100/\sqrt{25}} = 3.2$$

وتزيد قيمة t المحسوبة عن قيمة t الجدولية $t_{0.05} = 1.71$ بدرجات حرية 24 . وعليه فإننا نرفض H_0 ونقبل H_1 ، ويمكن للشركة أن تدعى بدرجة ثقة 95% أن صناديق الصابون من مصنع ١ تحتوى على أكثر من $1,000\,g$ من الصابون .

(ب)

$$H_0: \ \mu_1 = \mu_2 \quad \text{or} \quad H_0: \ \mu_1 - \mu_0 = 0$$

$$H_1: \ \mu_1 \neq \mu_2 \quad \text{or} \quad H_1: \ \mu_1 - \mu_0 \neq 0$$

$$\sigma_{\bar{X}_1 - \bar{X}_2} \cong \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{100^2}{25} + \frac{60^2}{25}} = \sqrt{544} \cong 23.32$$

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{X}_1 - \bar{X}_2}} = \frac{\bar{X}_1 - \bar{X}_2 - 0}{\sigma_{\bar{X}_1 - \bar{X}_2}} = \frac{1,064 - 1,024}{23.32} \cong 1.72$$

وهذا اختبار ذو ذيلين بدرجات حرية $n_1 + n_2 - 1 = 49$. وحيث أن القيمة الجدولية $2.00 < t_{0.025}$ بدرجات حرية 49 ، فتقبل الشركة بدرجة ثقة 95% الفرض أنه لايوجد اختلاف فى كية الصابون بالصناديق من إنتاج المصنعين .

الفصل السادس

تحليل الانحدار البسيط

٦–١ النموذج الخطى لمتغيرين

يستخدم النموذج الخطى ذى المتغيرين ، أو تحليل الانحدار البسيط ، لاختيار الفروض حول العلاقة بين متغير تابع ، Y ، ومتغير مستقل أو مفسر X ، وللتنبؤ . ويبدأ الانحدار الخطى البسيط عادة برسم مجموعة قيم XY فى شكل انتشار ثم التحديد بالنظر ما إذا كانت هناك علاقة خطية تقريبية .

$$Y_i = b_0 + b_1 X_i \qquad (٦ - ١)$$

وحيث أنه من غير المتوقع أن تقع النقاط تماماً على الخط ، فإن العلاقة الخطية التامة فى معادلة (٦ – ١) يجب أن تعدل لكى تضم حد تشويش عشوائى أو خطأ أى « عنصر عشوائى » ، u_i (أنظر قسم ١ – ٢ والمسألة ١ – ٨) :

$$Y_i = b_0 + b_1 X_i + u_i \qquad (٦ - ٢)$$

ويفترض فى حد الخطأ أنه (١) موزع طبيعياً ، (٢) وقيمته المتوقعة أى وسطه صفر ، (٣) وتباينه ثابت ، ويفترض أيضاً (٤) أن حدود الخطأ غير مترابطة بعضها ببعض (٥) وأن المتغير المفسر يأخذ قيماً ثابتة فى المعاينات المتكررة (حتى تكون X_i ، U_i غير مترابطة) .

مثال ١ : يعطى جدول ٦ – ١ « بوشلات » الحنطة للأكر (وحدة مساحة) ، Y الناتجة عن كميات مختلفة من السماد بالرطل ، X ، فى إحدى المزارع خلال 10 سنوات من ١٩٧١ إلى ١٩٨٠ . وهذه البيانات موضوعة فى شكل الانتشار المعطى فى الشكل ٦ – ١ . إن العلاقة بين X ، Y فى شكل ٦ – ١ تبدو خطية تقريباً (أى أن النقاط تقع على خط مستقيم أو بالقرب منه) .

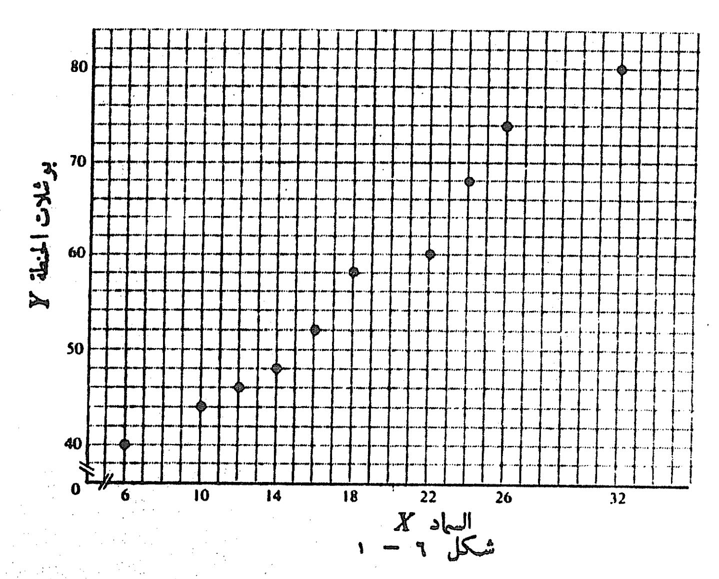
جدول ٦ – ١ الحنطة المنتجة مع السماد المستخدم

السنة	n	Y_i	X_i
1971	1	40	6
1972	2	44	10
1973	3	46	12
1974	4	48	14
1975	5	52	16
1976	6	58	18
1977	7	60	22
1978	8	68	24
1979	9	74	26
1980	10	80	32

٦–٢ طريقة المربعات الصغرى العادية

طريقة المربعات الصغرى العادية (OLS) هى أسلوب لتوفيق « أفضل » خط مستقيم لعينة مشاهدات XY . وهو يتضمن تصغير مجموع المربعات لانحرافات النقاط (الرأسية) عن الخط إلى أدنى حد ممكن :

$$\text{Min} \sum (Y_i - \hat{Y}_i)^2 \qquad (٦ - ٣)$$



شكل ٦ – ١

حيث تشير Y_i إلى المشاهدات الفعلية ، وتشير $\hat{Y}_i$ إلى القيم « الموفقة » المناظرة ، بحيث تكون $Y_i - \hat{Y}_i = e_i$ هى البواق .
ويعطى هذا الأسلوب المعادلتين الطبيعيتين التاليتين (أنظر مسألة ٦ – ٦) :

$$\sum Y_i = n\hat{b}_0 + \hat{b}_1 \sum X_i \qquad (٦ - ٤)$$

$$\sum X_i Y_i = \hat{b}_0 \sum X_i + \hat{b}_1 \sum X_i^2 \qquad (٦ - ٥)$$

حيث n عدد المشاهدات ، $\hat{b}_0$ ، $\hat{b}_1$ هى مقدرتان للمعلمتين الحقيقيتين b_0 ، b_1 .

وبحل المعادلتين (٦ – ٤) ، (٦ – ٥) آنياً ، نحصل على (أنظر المسألة ٦ – ٧ (أ)) :

$$\hat{b}_1 = \frac{n \sum X_i Y_i - \sum X_i \sum Y_i}{n \sum X_i^2 - \left(\sum X_i\right)^2} \qquad (٦ - ٦)$$

ونحصل على قيمة $\hat{b}_0$ كما يلى (أنظر المسألة ٦ – ٧ (ب))

$$\hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X} \qquad (٦ - ٧)$$

ومن المفيد عادة استخدام صيغة مكافئة لتقدير $\hat{b}_1$ (أنظر المسألة ٦ – ١٠ (أ)) :

$$\hat{b}_1 = \frac{\sum x_i y_i}{\sum x_i^2} \qquad (٦ - ٨)$$

حيث $x_i = X_i - \bar{X}$ ، $y_i = Y_i - \bar{Y}$ وتكون معادلة انحدار المربعات الصغرى المقدرة (OLS) :

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_i \qquad (٦ - ٩)$$

مثال ٢ – يوضح جدول ٦ – ٢ الحسابات اللازمة لتقدير معادلة الانحدار لمشكلة الحنطة والسماد فى جدول ٦ – ١ . باستخدام معادلة (٦ – ٨) ،

(أ) (ميل خط الانحدار المقدر)

$$\hat{b}_1 = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{956}{576} = 1.66$$

(ب) (تقاطع Y)

$$\hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X} \cong 57 - (1.66)(18) \cong 57 - 29.88 \cong 27.12$$

(ج) (معادلة الانحدار المقدرة)

$$\hat{Y}_i = 27.12 + 1.66 X_i$$

وعليه ، فعند $X_i = 0$ فإن $\hat{Y} = 27.12 = \hat{b}_0$. وعند $X_i = 18 = \bar{X}$ فإن $\hat{Y} = 27.12 + 16.66\,(18) = 57 = \bar{Y}$
من هذا يتضح أن خط الانحدار يمر خلال النقطة $\bar{X}\bar{Y}$ (أنظر شكل ٦ – ٢) .

جدول ٦ – ٢ الحنطة المنتجة مع الأسمدة المستخدمة : الحسابات

n	Y_i (الحنطة)	X_i (السماد)	y_i	x_i	$x_i y_i$	x_i^2
1	40	6	− 17	− 12	204	144
2	44	10	− 13	− 8	104	64
3	46	12	− 11	− 6	66	36
4	48	14	− 9	− 4	36	16
5	52	16	− 5	− 2	10	4
6	58	18	1	0	0	0
7	60	22	3	4	12	16
8	68	24	11	6	66	36
9	74	26	17	8	136	64
10	80	32	23	14	322	196
$n = 10$	$\sum Y_i = 570$ $\bar{Y} = 57$	$\sum X_i = 180$ $\bar{X} = 18$	$\sum y_i = 0$	$\sum x_i = 0$	$\sum x_i y_i = 956$	$\sum x_i^2 = 576$

٦–٣ اختبارات المعنوية لتقديرات المعالم

لاختيار المعنوية الإحصائية لتقديرات معامل الانحدار ، يلزمنا معرفة تباين $\hat{b}_0$ و $\hat{b}_1$ (أنظر مسألتى ٦ – ١٤ ، ٦ – ١٥) :

$$\operatorname{Var}\hat{b}_0 = \sigma_u^2 \frac{\sum X_i^2}{n\sum x_i^2} \qquad (٦ - ١٠)$$

$$\operatorname{Var}\hat{b}_1 = \sigma_u^2 \frac{1}{\sum x_i^2} \qquad (٦ - ١١)$$

وحيث أن σ_u^2 غير معلومة ، فإن تباين البواقى ، s^2 ، يستخدم كتقدير غير متحيز للتباين :

$$s^2 = \hat{\sigma}_u^2 = \frac{\sum e_i^2}{n-k} \qquad (٦ - ١٢)$$

حيث k عدد المعالم المقدرة .

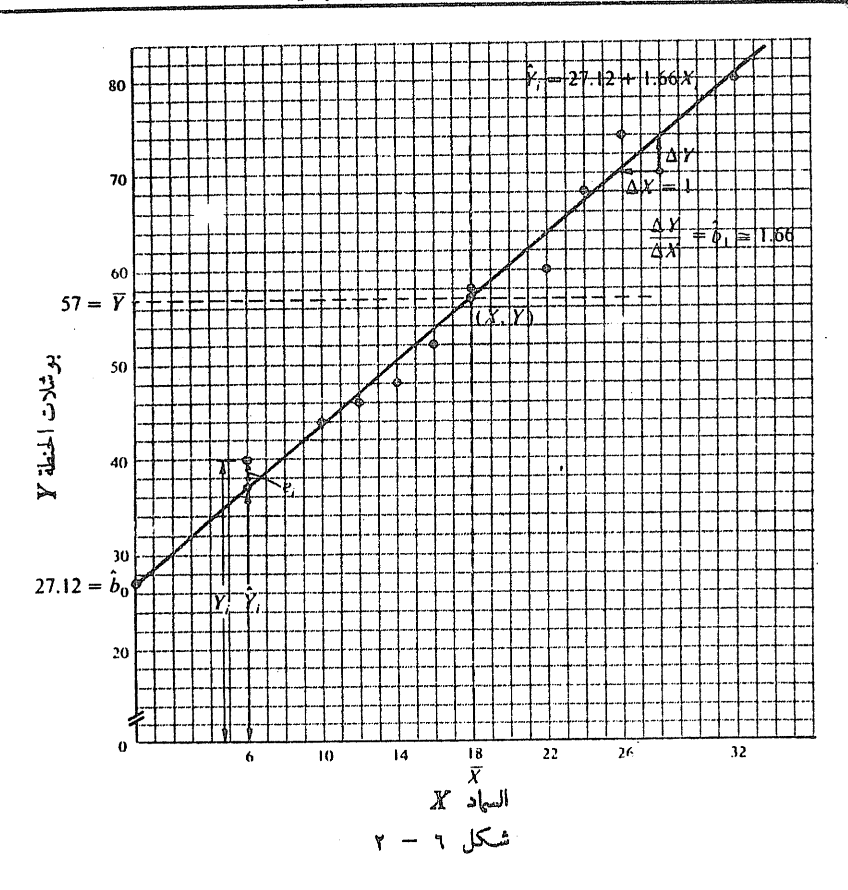
والمعادلات التالية تعطى تقديرات غير متحيزة لتباين $\hat{b}_0$ و $\hat{b}_1$

$$s_{\hat{b}_0}^2 = \frac{\sum e_i^2}{n-k}\,\frac{\sum X_i^2}{n\sum x_i^2} \qquad (٦ - ١٣)$$

$$s_{\hat{b}_1}^2 = \frac{\sum e_i^2}{n-k}\,\frac{1}{\sum x_i^2} \qquad (٦ - ١٤)$$

فتكون $s_{\hat{b}_0}$ و $s_{\hat{b}_1}$ هى الأخطاء المعيارية للتقدير . وحيث أن u موزعة طبيعياً فإن Y_i ، وبالتالى $\hat{b}_0$ و $\hat{b}_1$ تكون هى الأخرى موزعة طبيعياً ، ومن ثم يمكننا استخدام توزيع t بدرجات حرية $n-k$ ، لاختبار الفروض عن كل من $\hat{b}_0$ و $\hat{b}_1$ وعمل فترات ثقة لهما (أنظر قسمى ٤ – ٤ ، ٥ – ٢) .

مثال ٣ – جدول ٦ – ٣ (امتداد ٦ – ٢) يوضح الحسابات اللازمة لاختبار المعنوية الإحصائية لكل من $\hat{b}_0$ و $\hat{b}_1$. ولقد حصلنا على قيم $\hat{Y}_i$ فى جدول ٦ – ٣ بإحلال قيم X_i فى معادلة الانحدار المقدرة التى حصلنا عليها فى مثال ٢ .



شكل ٦ – ٢

(ونحصل على قيم y_i^2 بتربيع y_i من جدول ٦ – ٢ وسوف نستخدمها فى قسم ٦ – ٤) .

$$s^2_{\hat{b}_0} = \frac{\sum e_i^2}{n-k} \frac{\sum X_i^2}{n \sum x_i^2} \cong \frac{47.3056}{10-2} \frac{3,816}{10(576)} \cong 3.92 \quad \text{و} \quad s_{\hat{b}_0} = \sqrt{3.92} \cong 1.98$$

$$s^2_{\hat{b}_1} = \frac{\sum e_i^2}{(n-k)\sum x_i^2} \cong \frac{47.3056}{(10-2)576} \cong 0.01 \quad \text{و} \quad s_{\hat{b}_1} \cong \sqrt{0.01} \cong 0.1$$

جدول ٦ – ٣ حسابات الحنطة – السماد لاختبار معنوية المعالم

السنة	Y_i	X_i	$\hat{Y}_i$	e_i	e_i^2	X_i^2	x_i^2	y_i^2
1	40	6	37.08	2.92	8.5264	36	144	289
2	44	10	43.72	0.28	0.0784	100	64	169
3	46	12	47.04	− 1.04	1.0816	144	36	121
4	48	14	50.36	− 2.36	5.5696	196	16	81
5	52	16	53.68	− 1.68	2.8224	256	4	25
6	58	18	57.00	1.00	1.0000	324	0	1
7	60	22	63.64	− 3.64	13.2496	484	16	9
8	68	24	66.96	1.04	1.0816	576	36	121
9	74	26	70.28	3.72	13.8384	676	64	289
10	80	32	80.24	− 0.24	0.0576	1,024	196	529
$n = 10$				$\sum e_i = 0$	$\sum e_i^2 = 47.3056$	$\sum X_i^2 = 3,816$	$\sum x_i^2 = 576$	$\sum y_i^2 = 1,634$

وعليه فإن $t_1 = \frac{\hat{b}_1 - b_1}{s_{\hat{b}_1}} \cong \frac{1.66}{0.1} \cong 16.6$ و $t_0 = \frac{\hat{b}_0 - b_0}{s_{\hat{b}_0}} \cong \frac{27.12 - 0}{1.98} \cong 13.7$

وحيث أن كلا من t_0 و t_1 تتجاوز $t = 2.306$ بدرجات حرية 8 عند مستوى معنوية 5% (من ملحق ٥) ، نستنتج أن كلا من b_0 و b_1 معنوية إحصائياً بمستوى معنوية 5% .

٦ــ٤ اختبار جودة التوفيق والارتباط

كلما كانت المشاهدات أقرب إلى خط الانحدار (أى ، كلما صغرت البواقى) ، كلما زاد التغير فى Y الذى « تفسره » معادلة الانحدار المقدرة . والتغير الإجمالى فى Y يساوى التغير المفسر زائداً تغير البواقى :

$$\sum(Y_i - \bar{Y})^2 = \sum(\hat{Y}_i - \bar{Y})^2 + \sum(Y_i - \hat{Y}_i)^2 \qquad (٦ - ١٥)$$

التغير الإجمالى فى Y (أو إجمالى مجموع المربعات)		التغير المفسر فى Y (مجموع مربعات الانحدار)		تغير البواقى فى Y (مجموع مربعات الخطأ)
TSS	=	RSS	+	ESS

وبقسمة الطرفين على TSS نحصل على :

$$1 = \frac{\text{RSS}}{\text{TSS}} + \frac{\text{ESS}}{\text{TSS}}$$

ومن هنا يمكن تعريف معامل التحديد R^2 بأنه النسبة من التغير الإجمالى فى Y « الذى يفسره » انحدار Y على X :

$$R^2 = \frac{\text{RSS}}{\text{TSS}} = 1 - \frac{\text{ESS}}{\text{TSS}} \qquad (٦ - ١٦)$$

ويمكن حساب R^2 كالآتى :

$$R^2 = \frac{\sum \hat{y}_i^2}{\sum y_i^2} = 1 - \frac{\sum e_i^2}{\sum y_i^2} \qquad (٦ - ١٧)$$

حيث :

$$\sum \hat{y}_i^2 = \sum(\hat{Y}_i - \bar{Y}_i)^2$$

وتتراوح قيمة R^2 بين 0 (عندما لا تفسر معادلة الانحدار أياً من التغير فى Y) و 1 (عندما تقع كل النقاط على خط الانحدار) .

معامل الارتباط ، r ، يتم حسابه كالآتى (أنظر المسألة ٦ — ٢٢)

$$r = \sqrt{R^2} = \frac{\sum x_i y_i}{\sqrt{\sum x_i^2}\sqrt{\sum y_i^2}} = \sqrt{\hat{b}_1 \frac{\sum x_i y_i}{\sum y_i^2}} \qquad (٦ - ١٨)$$

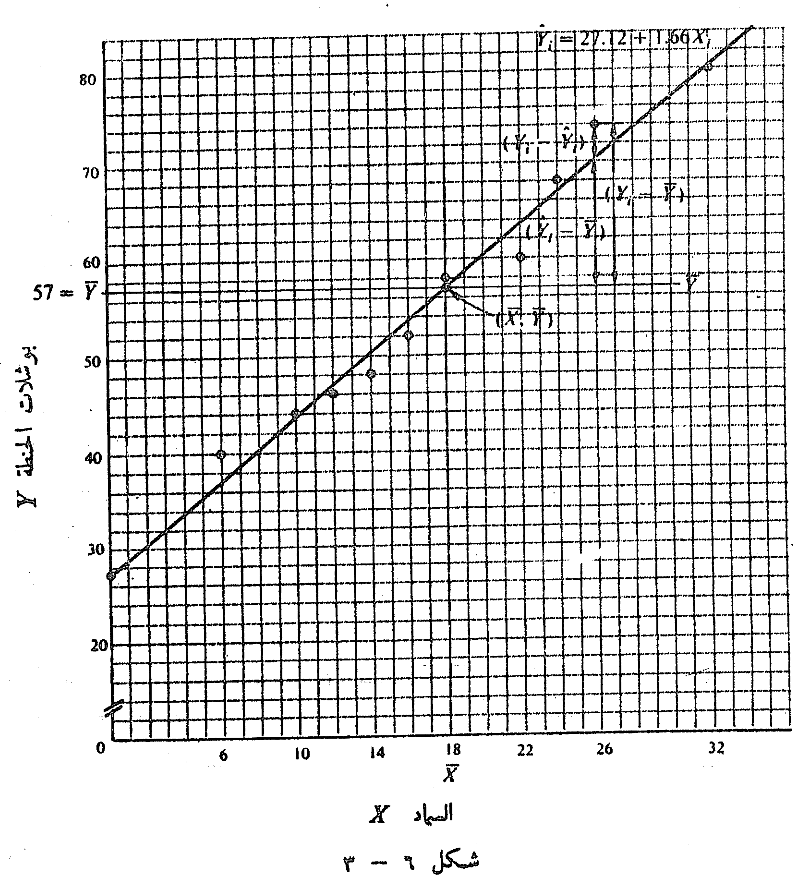
وتتراوح قيمة r من 1 — (للارتباط الخطى السالب التام) إلى 1 + (للارتباط الخطى الموجب التام) . **إن علاقة الارتباط بين متغيرين لا تعنى وجود** علاقة سببية أو علاقة تبعية بينهما . وفى حالة البيانات الكيفية ، يمكن استخدام **معامل ارتباط الرتب** ، r' (معامل سبيرمان) ، (أنظر المسألة ٦ — ٢٥) .

مثال ٤ — يمكن إيجاد معامل التحديد لمثال الحنطة — السماد من جدول ٦ — ٣ .

$$R^2 = 1 - \frac{\sum e_i^2}{\sum y_i^2} \cong 1 - \frac{47.31}{1,634} \cong 1 - 0.0290 \cong 0.9710, \text{ or } 97.10\%$$

أى أن معادلة الانحدار تفسر حوالى 97% من التغير الإجمالى فى إنتاج الحنطة . أما نسبة 3% الباقية فيمكن نسبتها إلى عوامل متضمنة فى حد الخطأ .

وتكون $r = \sqrt{R^2} \cong \sqrt{0.9710} \cong 0.9854$ أو 98.54% وهو موجب لأن $\hat{b}_1$ موجبة . ويوضح شكل ٦ – ٣ التغير الإجمالى ، والتغير المفسر ، وتغير البواقى فى Y .



شكل ٦ – ٣

٦–٥ خواص مقدرات طريقة المربعات الصغرى العادية

مقدرات المربعات الصغرى العادية (OLS) هى أفضل مقدرات خطية غير متحيزة (*BLUE*) . وعدم التحيز يعنى

$$E(\hat{b}) = b$$

بحيث أن

$$\text{Bias} = E(\hat{b}) - b$$

أما وصف مقدر بأنه « أفضل مقدر غير متحيز » أو أنه مقدر كفؤ فيعنى أنه ذو أصغر تباين . وبالتالى فإن مقدرات OLS هى الأفضل من بين كل المقدرات الخطية غير المتحيزة (أنظر المسألتين ٦ – ١٤ (أ) ، ٦ – ١٥ (أ)) . وتعرف هذه الخاصية بنظرية « جاوس – ماركوف » ، وهى تمثل أهم مبرر لاستخدام OLS .

أحياناً ، قد يرغب الباحث أن يقبل بعض التحيز فى مقابل تباين أصغر ، بتصغير متوسط مربع الخطأ ، MSE (أنظر المسألة ٦ – ٢٩) :

$$\text{MSE}(\hat{b}) = E(\hat{b} - b)^2 = \text{var}(\hat{b}) + (\text{bias}\,\hat{b})^2$$

ويكون المقدر متسقاً إذا اقتربت قيمته من المعلمة الحقيقية مع اقتراب حجم العينة من ما لا نهاية (بمعنى أنه غير متحيز فى اللانهاية) وينتهى توزيعه إلى المعلمة الحقيقية (أنظر المسألة ٦ – ٣٠) .

مثال ٥ – مقدرات OLS $\hat{b}_0$ و $\hat{b}_1$ السابق إيجادها فى مثال ٢ هى مقدرات خطية غير متحيزة لكل من b_0 و b_1 لأن

$$E(\hat{b}_0) = b_0 \quad \text{and} \quad E(\hat{b}_1) = b_1$$

كذلك فإن تباين $\hat{b}_0$ وتباين $\hat{b}_1$ السابق إيجادهما فى مثال ٣ أقل من تباين أى مقدرات خطية غير متحيزة أخرى . وعليه فإن كلا من $\hat{b}_0$ ، $\hat{b}_1$ BLUE

مسائل محلولة

النموذج الخطى ذو المتغيرين :

٦ – ١ ماذا يعنى وما هى وظيفة كل من (أ) تحليل الانحدار البسيط ؟ (ب) تحليل الانحدار الخطى ؟ (حـ) شكل الانتشار ؟ (د) حد الخطأ ؟

(أ) يستخدم « **الانحدار البسيط** » لاختبار الفروض عن العلاقة بين متغير تابع ، Y ، ومتغير مستقل أو مفسر ، X ، وللتنبؤ . قارن هذا بتحليل « **الانحدار المتعدد** » ، والذى ليس فيه متغير مستقل واحد وإنما اثنان أو أكثر من المتغيرات المستقلة أو المفسرة . وسوف نتناول الانحدار المتعدد فى الفصل السابع .

(ب) **تحليل الانحدار الخطى** يفترض أن هناك علاقة خطية تقريبية بين X و Y (بمعنى أن مجموعة قيم X ، Y فى العينة العشوائية تقع على أو قريباً من خط مستقيم) . قارن هذا بتحليل « **الانحدار غير الخطى** » (ويناقش فى قسم ٨ – ١) .

(حـ) **شكل الانتشار** هو شكل يعبر فيه عن كل زوج من المشاهدات المستقلة والتابعة بنقطة فى مستوى XY . والغرض منه هو أن نحدد (بالنظر) ما إذا كانت توجد علاقة خطية تقريبية بين المتغير التابع ، Y والمتغير المستقل أو المفسر .

(د) **حد الخطأ** (والمعروف أيضاً **بحد التشويش أو الحد العشوائى**) يقيس انحراف القيمة المشاهدة Y من خط الانحدار الحقيقى (غير المشاهد) . وتنشأ حدود الخطأ هذه والتى يدل عليها u_i ، s بسبب (١) وجود عدة متغيرات مفسرة ذات تأثير ضئيل أو غير منتظم على Y وقد استبعدت من العلاقة الخطية التامة فى معادلة ٦ – ١ ، (٢) أخطاء ممكنة فى قياس Y . (٣) السلوك الإنسانى العشوائى (أنظر المسألة ١ – ٨) .

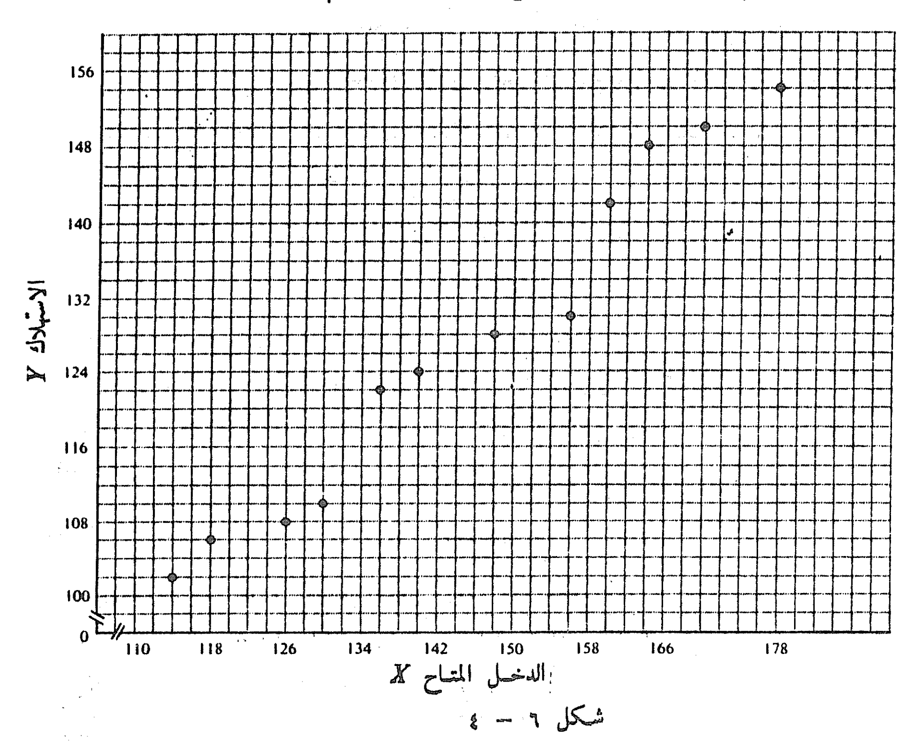
٦ – ٢ ارسم شكل انتشار لبيانات جدول ٦ – ٤ وحدد بالنظر ما إذا كانت توجد علاقة خطية تقريبية بين Y (إجمالى الإنفاق الاستهلاكى ، ببلايين الدولارات الأمريكية) ، X (إجمالى الدخل المتاح ، ببلايين الدولارات الأمريكية أيضاً) لإثنى عشر عاماً من ١٩٧١ إلى ١٨٩٢ .

جدول ٦ – ٤ إجمالى الاستهلاك والدخل المتاح

السنة	n	Y_i	X_i
1971	1	102	114
1972	2	106	118
1973	3	108	126
1974	4	110	130
1975	5	122	136
1976	6	124	140
1977	7	128	148
1978	8	130	156
1979	9	142	160
1980	10	148	164
1981	11	150	170
1982	12	154	178

من شكل ٦ – ٤ يمكن ملاحظة أن العلاقة بين الإنفاق الاستهلاكى ، Y ، والدخل المتاح ، X ، هى تقريباً خطية ، كما يقضى نموذج الانحدار الخطى .

٦ – ٣ عبر عن العلاقة العامة بين الاستهلاك ، Y ، والدخل المتاح ، X ، فى شكل (أ) صورة خطية تامة (ب) صورة عشوائية (جـ) لماذا تتوقع أن معظم القيم المشاهدة للمتغير Y لا تقع تماماً على خط مستقيم ؟



شكل ٦ – ٤

(أ) العلاقة العامة المحددة أو التامة بين إجمالى الإنفاق الاستهلاكى ، Y ، وإجمالى الدخل المتاح X ، يمكن كتابتها كالآتى :

$$Y_i = b_0 + b_1X_i \qquad (٦ - ١)$$

حيث تشير i إلى سنة ما فى تحليل السلاسل الزمنية (كما فى البيانات فى جدول ٦ – ٤) أو إلى كل وحدة اقتصادية (مثل الأسرة) فى تحليل البيانات المقطعية . وفى المعادلة (٦ – ١) b_0 ، b_1 هى ثوابت غير معلومة تسمى معالم . المعلمة b_0 هى الثابت أو الجزء المقطوع من محور Y ، بينما b_1 تقيس $\Delta Y/\Delta X$ ، والتى تشير ، بالنسبة للمسألة ٦ – ٢ ، إلى الميل الحدى للاستهلاك ، MPC (أنظر قسم ١ – ٢) . أما العلاقة الخطية المعينة المناظرة للعلاقة الخطية العامة فى معادلة (٦ – ١) فنحصل عليها بتقدير قيم b_0 ، b_1 (ويمثلها $\hat{b}_0$ ، $\hat{b}_1$ ، وتقرأ b "sub one hat"، b " sub zero hat" .

(ب) يمكن جعل العلاقة الخطية التامة فى معادلة (٦ – ١) عشوائية باضافة حد « تشويش » عشوائى أى حد خطأ ، u_i ، فتصبح

$$Y_i = b_0 + b_1X_i + u_i \qquad (٦ - ٢)$$

(جـ) لا يتوقع أن تقع معظم قيم Y المشاهدة تماماً على خط مستقيم (١) لأنه حتى لو سلمنا أن الاستهلاك Y ، يعتمد أساساً على الدخل المتاح X ، فإنه أيضاً من الممكن أن يعتمد على العديد من المتغيرات الأخرى المحذوفة ذات التأثير الصغير

أو غير المنتظم على Y (لو كان لبعض هذه المتغيرات الأخرى تأثير جوهرى أو منتظم على Y ، لكانت قد دخلت فى العلاقة كتغيرات مفسرة إضافية ، كما فى نموذج الانحدار المتعدد) (٢) بسبب الأخطاء الممكنة فى قياس Y (٣) بسبب السلوك الإنسانى العشوائى المتأصل ، والذى يؤدى عادة إلى قيم مختلفة للمتغير Y مقابل نفس القيمة للمتغير X فى ظل ظروف متطابقة (أنظر المسألة ١ – ٨) .

٦ – ٤ أذكر الفروض الخمسة لنموذج الانحدار الخطى الكلاسيكى (OLS) واعط تفسيرا بديهياً لمعنى كل منها والحاجة إليه .

١ – الفرض الأول لنموذج الانحدار الخطى الكلاسيكى (OLS) هو أن حد الخطأ العشوائى u يتبع التوزيع الطبيعى . وكنتيجة فان Y وتوزيع المعاينة لمعالم الانحدار تتبع أيضاً التوزيع الطبيعى ، بحيث يمكن القيام باختبارات لمعنوية هذه المعالم (أنظر أقسام ٤ – ٢ ، ٥ – ٢ ، ٦ – ٣) .

٢ – والفرض الثانى هو أن القيمة المتوقعة لحد الخطأ أى وسطه يساوى الصفر أى

$$E(u_i)=0 \qquad (١٩-٦)$$

وبسبب هذا الفرض فان معادلة (٦ – ١) تعطى متوسط قيمة Y . وبالتحديد ، حيث أنه يفترض أن X ثابتة ، فإن قيمة Y فى معادلة (٦ – ٢) تتغير فوق أو تحت وسطها مع زيادة u أو نقصها عن 0 . وحيث أننا نفترض أن متوسط u يساوى 0 ، فإن معادلة (٦ – ١) تعطى القيمة المتوسطة للمتغير Y .

٣ – الفرض الثالث هو أن تباين حد الخطأ ثابت فى كل فترة ولكل قيم X ، أى

$$E(u_i)^2=\sigma_u^2 \qquad (٢٠-٦)$$

ويكفل هذا الفرض أن كل مشاهدة يمكن الاعتماد عليها بنفس القدر ، بحيث تكون تقديرات معاملات الانحدار كفؤا ، وتكون اختبارات الفروض الخاصة بها غير متحيزة . ويمكن تلخيص الفروض الثلاثة الأولى على حد الخطأ كالآتى

$$u\sim N(0,\sigma_u^2) \qquad (٢١-٦)$$

٤ – الفرض الرابع هو أن القيمة التى يأخذها حد الخطأ فى فترة ما تكون غير مرتبطة أو غير متعلقة بقيمته فى أى فترة أخرى ، أى

$$E(u_iu_j)=0 \qquad \text{for } i\neq j;\quad i,j=1,2,\ldots,n \qquad (٢٢-٦)$$

وهذا يكفل أن القيمة المتوسطة للمتغير Y تعتمد فقط على X وليس على u ، وهذا ، مرة أخرى ، مطلوب للحصول على تقديرات كفء لمعاملات الانحدار واختبارات غير متحيزة لمعنويتها .

٥ – الفرض الخامس أن المتغير المفسر يأخذ قيما ثابتة والتى يمكن الحصول عليها فى العينات المتكررة ، بحيث أن المتغير المفسر يكون هو الآخر غير مرتبط بعنصر الخطأ ، أى

$$E(X_iu_i)=0 \qquad (٢٣-٦)$$

ويوضح هذا الفرض لتبسيط التحليل .

طريقة المربعات الصغرى العادية :

٦ – ٥ (أ) يقصد بطريقة المربعات الصغرى العادية أو ، *OLS* ، لتقدير « أفضل خط مستقيم يوفق عينة المشاهدات X ، Y ؟

(ب) لماذا نأخذ الانحرافات الرأسية ؟ (ج) لماذا لا نأخذ ببساطة مجموع الانحرافات دون تربيعها؟ (د) لماذا لا نأخذ مجموع الانحرافات المطلقة ؟

(أ) تعطى طريقة OLS أفضل خط مستقيم يوفق مشاهدات العينة XY بمعنى أنها تعطى أقل مجموع مربعات (الرأسية) لانحرافات كل مشاهدة عن الخط المستقيم فى الرسم .

(ب) نأخذ الانحرافات الرأسية لأننا نحاول أن نفسر وأن نتنبأ بالتغيرات فى Y التى تقاس على المحور الرأسى .

(جـ) لا يمكننا أخذ مجموع الانحرافات لكل مشاهدة عن خط OLS لأن الانحرافات المتساوية فى الحجم والمختلفة فى الإشارة يلغى بعضها البعض ، فيكون مجموع الانحرافات مساوياً للصفر (أنظر جدول ٦ – ٢) .

(د) بأخذ مجموع الانحرافات المطلقة نتجنب مشكلة أن يصبح مجموع الانحرافات 0 . ولكن ، يفضل مجموع مربعات الانحرافات لمعاقبة الانحرافات الكبيرة أكثر من الانحرافات الصغيرة .

٦ – ٦ بدءاً من معادلة ٦ – ٣ والتى تدعو إلى تصغير مجموع مربعات الانحرافات أو البواق ، اشتق (أ) المعادلة الطبيعية (٦ – ٤) (ب) المعادلة الطبيعية (٦ – ٥) . (القارئ غير الملم بالتفاضل يمكنه أن يتخطى هذه المسألة) .

(أ) $$\sum e_i^2 = \sum (Y_i - \hat{Y}_i)^2 = \sum (Y_i - \hat{b}_0 - \hat{b}_1 X_i)^2$$

تشتق المعادلة الطبيعية (٦ – ٤) بتصغير $\sum e_i^2$ بالنسبة إلى $\hat{b}_0$.

$$\frac{\partial \sum e_i^2}{\partial \hat{b}_0} = \frac{\partial \sum (Y_i - \hat{b}_0 - \hat{b}_1 X_i)^2}{\partial \hat{b}_0} = 0$$

$$2\sum (Y_i - \hat{b}_0 - \hat{b}_1 X_i)(-1) = 0$$

$$\sum (Y_i - \hat{b}_0 - \hat{b}_1 X_i) = 0$$

$$\sum Y_i = n\hat{b}_0 + \hat{b}_1 \sum X_i \qquad (٦ – ٤)$$

(ب) وتشتق المعادلة الطبيعية (٦ – ٥) بتصغير $\sum e_i^2$ بالنسبة إلى $\hat{b}_1$:

$$\frac{\partial \sum e_i^2}{\partial \hat{b}_1} = \frac{\partial \sum (Y_i - \hat{b}_0 - \hat{b}_1 X_i)^2}{\partial \hat{b}_1} = 0$$

$$2\sum (Y_i - \hat{b}_0 - \hat{b}_1 X_i)(-X_i) = 0$$

$$\sum (Y_i X_i - \hat{b}_0 X_i - \hat{b}_1 X_i^2) = 0$$

$$\sum Y_i X_i = \hat{b}_0 \sum X_i + \hat{b}_1 \sum X_i^2 \qquad (٦ – ٥)$$

٦ – ٧ اشتق (أ) معادلة (٦ – ٦) لإيجاد $\hat{b}_1$ ، (ب) معادلة (٦ – ٧) لإيجاد $\hat{b}_0$ (ارشاد بالنسبة لجزء (أ) : ابدأ بضرب المعادلة (٦ – ٥) فى n والمعادلة (٦ – ٤) فى ΣX_i) .

(أ) بضرب معادلة (٦ – ٥) فى n ومعادلة (٦ – ٤) فى ΣX_i ، نحصل على :

$$n\sum X_i Y_i = \hat{b}_0 n \sum X_i + \hat{b}_1 n \sum X_i^2 \qquad (٦ – ٢٤)$$

$$\sum X_i \sum Y_i = \hat{b}_0 n \sum X_i + \hat{b}_1 \left(\sum X_i\right)^2 \qquad (٦ – ٢٥)$$

بطرح معادلة (٦ – ٢٥) من معادلة (٦ – ٢٤) نحصل على :

$$n\sum X_i Y_i - \sum X_i \sum Y_i = \hat{b}_1 \left[n\sum X_i^2 - \left(\sum X_i\right)^2 \right] \qquad (٦ – ٢٦)$$

بحل معادلة (٦ – ٢٦) بالنسبة إلى $\hat{b}_1$ ، نحصل على

$$\hat{b}_1 = \frac{n\sum X_iY_i - \sum X_i \sum Y_i}{n\sum X_i^2 - \left(\sum X_i\right)^2} \qquad (٦ - ٦)$$

(ب) يمكن الحصول على معادلة ٦ – ٧ بحل معادلة (٦ – ٤) فى $\hat{b}_0$:

$$\sum Y_i = n\hat{b}_0 + \hat{b}_1 \sum X_i \qquad (٦ - ٤)$$

$$\hat{b}_0 = \frac{\sum Y_i}{n} - \hat{b}_1 \frac{\sum X_i}{n}$$

$$= \bar{Y} - \hat{b}_1 \bar{X} \qquad (٦ - ٧)$$

٦ – ٨ (أ) أذكر الفرق بين b_0 ، b_1 ، من ناحية $\hat{b}_0$ ، $\hat{b}_1$ من ناحية أخرى . (ب) أذكر الفرق بين u_i ، e_i (جـ) أكتب معادلات العلاقات الحقيقية والمقدرة بين X ، Y (د) أكتب المعادلات الحقيقية والمقدرة لخطوط الانحدار بين X ، Y .

(أ) b_0 ، b_1 هى معالم خط الانحدار الحقيقى ولكن غير المعلوم ، بينما $\hat{b}_0$ ، $\hat{b}_1$ هى معالم خط الانحدار المقدر .

(ب) u_i هى حد تشويش عشوائى ، أو الخطأ فى العلاقة الحقيقية غير المعلومة بين X ، Y . بينما e_i هى البواقى بين قيمة كل مشاهدة Y والقيمة المناظرة لها $\hat{Y}$ ، فى العلاقة المقدرة .

(جـ) المعادلات للعلاقات الحقيقية والمقدرة بين X ، Y هى على الترتيب ،

$$Y_i = b_0 + b_1X_i + u_i \qquad (٦ - ٢)$$

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1X_i + e_i \qquad (٦ - ٢٧)$$

(د) المعادلات للانحدار الحقيقى والمقدرة بين X ، Y هى على الترتيب ،

$$E(Y_i) = b_0 + b_1X_i \qquad (٦ - ٢٨)$$

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1X_i \qquad (٦ - ٩)$$

٦ – ٩ أوجد معادلة انحدار جدول الاستهلاك فى جدول (٦ – ٤) ، باستخدام معادلة (٦ – ٦) لإيجاد $\hat{b}_1$ (ب) ارسم خط الانحدار وبين انحرافات كل Y_i عن القيمة المناظرة $\hat{Y}_i$.

(أ) يوضح جدول ٦ – ٥ حسابات إيجاد $\hat{b}_1$ ، $\hat{b}_0$ لبيانات جدول ٦ – ٤ .

$$\hat{b}_1 = \frac{n\sum X_iY_i - \sum X_i \sum Y_i}{n\sum X_i^2 - \left(\sum X_i\right)^2} = \frac{(12)(225,124) - (1,740)(1,524)}{(12)(257,112) - (1,740)^2} = \frac{2,701,488 - 2,651,760}{3,085,344 - 3,027,600} = \frac{49,728}{57,744} \cong 0.86$$

$$\hat{b}_0 = \bar{Y} - \hat{b}_1\bar{X} \cong 127 - 0.86(145) \cong 127 - 124.70 \cong 2.30$$

وتكون المعادلة لانحدار الاستهلاك المقدر

$$\hat{Y}_i = 2.30 + 0.86X_i$$

جدول ٦ – ٥ الاستهلاك الإجمالي والدخل المتاح : حسابات

n	Y_i	X_i	$X_i Y_i$	X_i^2
1	102	114	11,628	12,996
2	106	118	12,508	13,924
3	108	126	13,608	15,876
4	110	130	14,300	16,900
5	122	136	16,592	18,496
6	124	140	17,360	19,600
7	128	148	18,944	21,904
8	130	156	20,280	24,336
9	142	160	22,720	25,600
10	148	164	24,272	26,896
11	150	170	25,500	28,900
12	154	178	27,412	31,684
$n = 12$	$\sum Y_i = 1,524$ $\bar{Y} = 127$	$\sum X_i = 1,740$ $\bar{X} = 145$	$\sum X_i Y_i = 225,124$	$\sum X_i^2 = 257,112$

(ب) لرسم معادلة الانحدار ، نحتاج إلى نقطتين على خط الانحدار . فمثلا ،

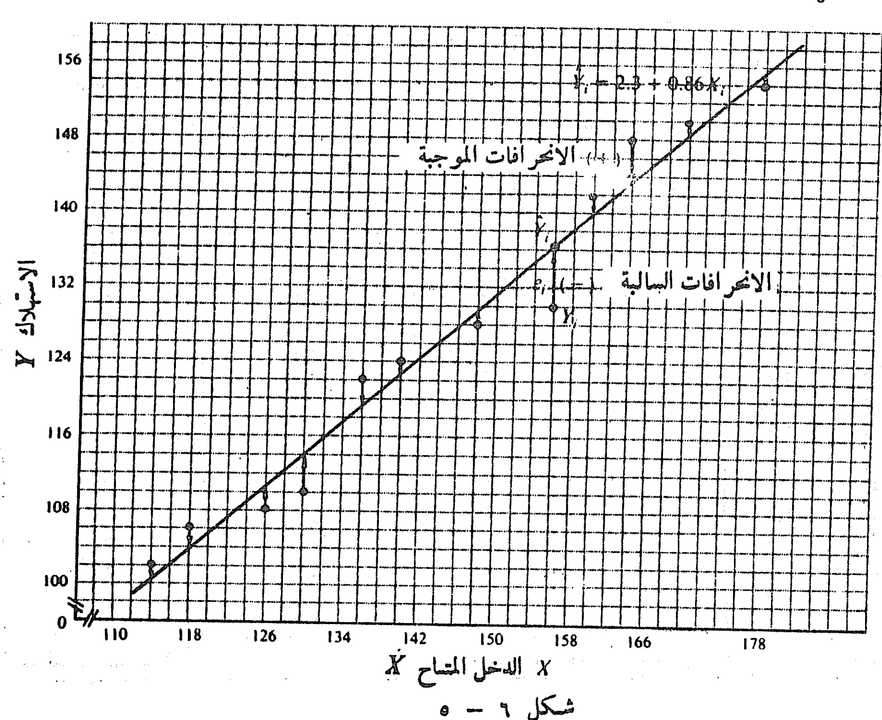
عند $X_i = 114$ تكون $\hat{Y}_i = 2.30 + 0.86 \ (114) = 100.34$ ،

وعند $X_i = 178$ تكون $\hat{Y}_i = 2.30 + 0.86 \ (178) = 155.38$

وقد رسم خط الانحدار الاستهلاك فى شكل ٦ – ٥ ، حيث توضح أيضاً البواقى السالبة والموجبة . ويمثل خط الانحدار أفضل توفيق للعينة العشوائية لمشاهدات الاستهلاك – الدخل المتاح بمعنى أنه يجعل مجموع مربعات الانحرافات (الرأسية) عن الخط أقل ما يمكن .

٦ – ١٠ (ا) بدءاً بالمعادلة (٦ – ٦) ، اشتق معادلة للمعامل $\hat{b}_1$ باستخدام الانحرافات عن المتوسطات عندما يكون $\bar{X} = \bar{Y} = 0$

(ب) ما هى قيمة $\hat{b}_0$ عند $\bar{X} = \bar{Y} = 0$. ؟



شكل ٦ – ٥

(أ) بدءاً بمعادلة $\hat{b}_1$ (٦ – ٦) :

$$\hat{b}_1 = \frac{n\sum X_i Y_i - \sum X_i \sum Y_i}{n\sum X_i^2 - \left(\sum X_i\right)^2} \qquad (٦ - ٦)$$

بقسمة البسط والمقام على n^2 نحصل على

$$\hat{b}_1 = \frac{\sum X_i Y_i/n - \left(\sum X_i/n\right)\left(\sum Y_i/n\right)}{\sum X_i^2/n - \left(\sum X_i/n\right)^2}$$

$$= \frac{\sum X_i Y_i/n - \bar{X}\bar{Y}}{\sum X_i^2/n - \bar{X}^2}$$

وبحذف n's من البسط والمقام حيث $\bar{X} = \bar{Y} = 0$

$$= \frac{\sum X_i Y_i}{\sum X_i^2}$$

$$= \frac{\sum x_i y_i}{\sum x_i^2} \quad \text{حيث } \bar{X} = \bar{Y} = 0 \qquad (٦ - ٨)$$

(ب) بدءاً بمعادلة $\hat{b}_0$ (٦ – ٧) :

$$\hat{b}_0 = \bar{Y} - \hat{b}_1\bar{X} \qquad (٦ - ٧)$$

وباحلال 0 محل كل من $\bar{X}$ ، $\bar{Y}$ نحصل على :

$$\hat{b}_0 = 0 - \hat{b}_1(0) = 0$$

٦ – ١١ بالنسبة لبيانات جدول ٦ – ٤ (أ) أوجد قيمة $\hat{b}_1$ باستخدام معادلة (٦ – ٨) ، (ب) ارسم خط الانحدار مع قياس المتغيرات كانحرافات عن متوسطاتها . كيف يقارن خط الانحدار هذا بخط الانحدار فى شكل ٦ – ٥ ؟

(أ) جدول ٦ – ٦ يوضح الحسابات لإيجاد $\hat{b}_1$ لبيانات جدول ٦ – ٤ .

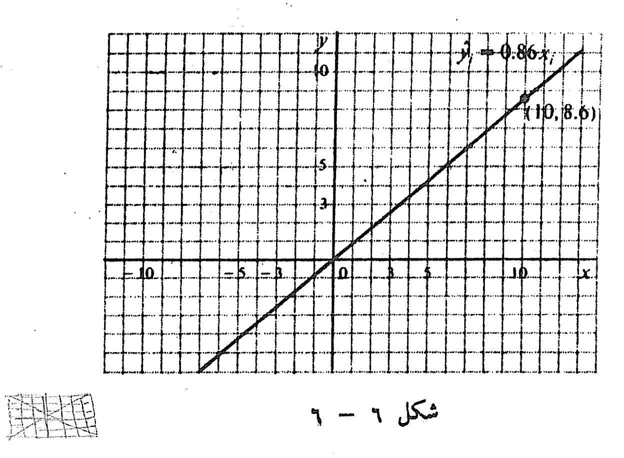
جدول ٦ – ٦ الاستهلاك الإجمالى والدخل المتاح : حسابات بديلة

n	Y_i	X_i	y_i	x_i	$x_i y_i$	x_i^2
1	102	114	– 25	– 31	775	961
2	106	118	– 21	– 27	567	729
3	108	126	– 19	– 19	361	361
4	110	130	– 17	– 15	255	225
5	122	136	–5	–9	45	81
6	124	140	–3	–5	15	25
7	128	148	1	3	3	9
8	130	156	3	11	33	121
9	142	160	15	15	225	225
10	148	164	21	19	399	361
11	150	170	23	25	575	625
12	154	178	27	33	891	1,089
			$\sum y_i = 0$	$\sum x_i = 0$	$\sum x_i y_i = 4{,}144$	$\sum x_i^2 = 4{,}812$

فى صورة انحرافات (لاحظ أن $\Sigma y_i = \Sigma x_i = 0$) :

$$\hat{b}_1 = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{4,144}{4,812} \cong 0.86$$ (نفس القيمة كما فى المسألة ٦ – ٩ (أ))

(ب) نعلم من المسألة ٦ – ١٠ (ب) أن خط الانحدار يمر بنقطة الأصل عندما يرسم فى شكل بيانى محاوره تقيس المتغيرات فى صورة انحرافات عن متوسطاتها ، ومن (أ) من هذه المسألة نعلم أن خط الانحدار هذا له نفس ميل خط الانحدار فى شكل ٦ – ٥ . أنظر شكل ٦ – ٦ .



شكل ٦ – ٦

٦ – ١٢ فى سياق المسألة ٦ – ٩ (أ) ما معنى كل من (أ) المقدر $\hat{b}_0$ ؟ (ب) المقدر $\hat{b}_1$ ؟ (جـ) أوجد المرونة الدخلية للاستهلاك .

(أ) المقدر $\hat{b}_0 \cong 2.30$ هو الجزء المقطوع من Y ، أو قيمة الاستهلاك الإجمالى بالبليون دولار ، عندما يساوى الدخل المتاح ، أيضاً بالبليون دولار ، صفرا . وحقيقة أن $\hat{b}_0 > 0$ يؤكد ما كان متوقعاً على أسس نظرية فى مثال ٣ فى الفصل الأول .

(ب) المقدر $\hat{b}_1 = dY/dX \cong 0.86$ هو ميل خط الانحدار المقدر . وهو يقيس الميل الحدى للاستهلاك ، MPC ، أو التغير فى الاستهلاك لوحدة التغير فى الدخل . مرة أخرى ، حقيقة أن $0 < \hat{b}_1 < 1$ يؤكد ما كان متوقعاً على أسس نظرية فى مثال ٣ فى الفصل الأول .

(جـ) المرونة الدخلية للاستهلاك η تقيس نسبة التغير فى الاستهلاك الناتجة عن تغير نسبى معين فى الدخل المتاح . وحيث أن المرونة تتغير عادة عند كل نقطة للدالة ، فإنها تقاس عند المتوسطات

$$\eta = \hat{b}_1 \frac{\bar{X}}{\bar{Y}} \qquad (٦ - ٢٩)$$

من بيانات جدول ٦ – ٤ .

$$\eta = 0.86 \frac{145}{127} \cong 0.98$$

لاحظ أن المرونة ، بالمقارنة بالميل ، هى رقم مطلق (بدون تمييز) .

اختبارات المعنوية لتقديرات المعالم :

٦ – ١٣ عرف (أ) σ_u^2 و S^2 (ب) تباين $\hat{b}_0$ وتباين $\hat{b}_1$ (جـ) $s_{\hat{b}_0}^2$ و $s_{\hat{b}_1}^2$ و (د) $s_{\hat{b}_0}$ ، و $s_{\hat{b}_1}$.

(أ) $\sigma_u{}^2$ هو تباين حد الخطأ فى العلاقة الحقيقية بين X_i و Y_i . ولكن $s^2 = \hat{\sigma}_u^2 = \sum e_i^2/(n-k)$ وهو تباين البواق وهو تقدير (غير متحيز) للتباين σ_u^2 غير المعلوم . K هى عدد المعالم المقدرة . فى تحليل الانحدار البسيط ، $k = 2$ وعليه فان $n - k = n - 2$ وتشير إلى درجات الحرية .

(ب) تباين $\hat{b}_0 = \sigma_u^2 \sum X_i/n\sum x_i^2$ بينما تباين $\hat{b}_1 = \sigma_u^2/\sum x_i^2$. إن تباينى $\hat{b}_0$ و $\hat{b}_1$ (أو تقديراتهما) مطلوبة لاختبار الفروض عن وتكوين فترات الثقة لكل من $\hat{b}_0$ و $\hat{b}_1$.

(جـ) $$s_{\hat{b}_1}^2 = \frac{s^2}{\sum x_i^2} = \frac{\sum e_i^2}{(n-k)\sum x_i^2} \quad \text{و} \quad s_{\hat{b}_0}^2 = s^2\frac{\sum X_i^2}{n\sum x_i^2} = \frac{\sum e_i^2 \sum X_i^2}{(n-k)n\sum x_i^2}$$

$s_{\hat{b}_0}^2$ ، $s_{\hat{b}_1}^2$ هى ، على الترتيب ، تقديرات (غير متحيزة) لتباين $\hat{b}_0$ وتباين $\hat{b}_1$ ، غير المعلومين نظراً لأن σ_u^2 غير معلومة .

(د) $s_{\hat{b}_0} = \sqrt{s_{\hat{b}_0}^2}$ و $s_{\hat{b}_1} = \sqrt{s_{\hat{b}_1}^2}$. $s_{\hat{b}_0}^2$ و $s_{\hat{b}_1}$ هى على الترتيب ، الانحرافات المعيارية $\hat{b}_0$ و $\hat{b}_1$ وتسمى الأخطاء المعيارية .

٦ - ١٤ أثبت أن (أ) متوسط $\hat{b}_1 = b_1$ ، و (ب) تباين $\hat{b}_1 = \sigma_u^2/\sum x_i^2$ (جـ) متوسط $\hat{b}_0 = b_0$ و (د) تباين $\hat{b}_0 = \sigma_u^2(\sum X_i^2/n\sum x_i^2)$

$$\hat{b}_1 = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{\sum x_i(Y_i - \bar{Y})}{\sum x_i^2} = \frac{\sum x_i Y_i}{\sum x_i^2} - \frac{\sum x_i \bar{Y}}{\sum x_i^2} = \frac{\sum x_i Y_i}{\sum x_i^2} \quad \text{حيث} \quad \sum x_i = 0$$

$$\hat{b}_1 = \frac{\sum x_i Y_i}{\sum x_i^2} = \sum c_i Y_i$$

حيث $c_i = x_i/\sum x_i^2 =$ ثابت طبقاً لفرض 5 (قسم ٦ - ١)

$$\hat{b}_1 = \sum c_i Y_i = \sum c_i(b_0 + b_1 X_i + u_i) = b_0\sum c_i + b_1\sum c_i X_i + \sum c_i u_i = b_1 + \sum c_i u_i = b_1 + \frac{\sum x_i u_i}{\sum x_i^2}$$

حيث $\sum c_i = \sum x_i/\sum x_i^2 = 0$ (لأن $\sum x_i = 0$) و

$$\sum c_i X_i = \frac{\sum x_i X_i}{\sum x_i^2} = \frac{\sum (X_i - \bar{X})X_i}{\sum (X_i - \bar{X})^2} = \frac{\sum X_i^2 - \bar{X}\sum X_i}{\sum X_i^2 - 2\bar{X}\sum X_i + n\bar{X}^2} = \frac{\sum X_i^2 - \bar{X}\sum X_i}{\sum X_i^2 - \bar{X}\sum X_i} = 1$$

$$E(\hat{b}_1) = E(b_1) + E\left[\frac{\sum x_i u_i}{\sum x_i^2}\right] = E(b_1) + \frac{1}{\sum x_i^2}E(\sum x_i u_i) = b_1 \quad \text{وسط} \quad b_1$$

وحيث أن b_1 ثابت ، $(\Sigma x_i u_i) = 0$ طبقاً لفرض 5 (قسم ٦ - ١) .

(ب) من (أ)

$$\hat{b}_1 = \frac{\sum x_i Y_i}{\sum x_i^2} = \sum c_i Y_i$$

$$\text{Var}\,\hat{b}_1 = \text{var}\left(\sum c_i Y_i\right)^2 = \sum c_i^2 \text{var}\, Y_i = \sum c_i^2 \sigma_u^2$$

وحيث أن Y_i تتغير بسبب u_i فقط ثبات X_i فرضاً .

$$\text{Var}\,\hat{b}_1 = \sum c_i^2\sigma_u^2 = \sum\left(\frac{x_i}{\sum x_i^2}\right)^2\sigma_u^2 = \frac{\sum x_i^2}{\left(\sum x_i^2\right)^2}\sigma_u^2 = \frac{\sigma_u^2}{\sum x_i^2}$$

(جـ)

$$\hat{b}_0 = \bar{Y} - \hat{b}_1\bar{X} = \frac{\sum Y_i}{n} - \bar{X}\sum c_i Y_i$$

$$\hat{b}_0 = \frac{\sum Y_i}{n} - \bar{X}\sum c_i Y_i = \sum\left(\frac{1}{n} - \bar{X}c_i\right)Y_i$$

(٣٢) (من معادلة (٦ - ١) فى المسألة ٦ - ٨ (د))

$$E(b_0) = \sum\left(\frac{1}{n} - \bar{X}c_i\right)E(Y_i) = \sum\left(\frac{1}{n} - \bar{X}c_i\right)(b_0 + b_1X_i)$$

وبفك الأقواس

$$E(b_0) = \sum\left(\frac{b_0}{n} - \bar{X}c_ib_0 + \frac{b_1X_i}{n} - \bar{X}c_ib_1X_i\right) = b_0 + b_1\bar{X} - b_1\bar{X} = b_0$$

إذ أن $\sum c_i = 0$ و $\sum c_iX_i = 1$ ، من (أ)

(د) رأينا فى (جـ) أن

$$\hat{b}_0 = \sum\left(\frac{1}{n} - \bar{X}c_i\right)Y_i$$

$$\operatorname{Var}\hat{b}_0 = \operatorname{var}\left[\sum\left(\frac{1}{n} - \bar{X}c_i\right)Y_i\right] = \sum\left(\frac{1}{n} - \bar{X}c_i\right)^2 \operatorname{var} Y_i = \sigma_u^2\sum\left(\frac{1}{n} - \bar{X}c_i\right)^2$$

$$\operatorname{Var}\hat{b}_0 = \sigma_u^2\sum\left(\frac{1}{n^2} - \frac{2\bar{X}c_i}{n} + \bar{X}^2c_i^2\right) = \sigma_u^2\left(\frac{1}{n} + \frac{\bar{X}^2}{\sum x_i^2}\right) = \sigma_u^2\frac{\sum x_i^2 + n\bar{X}^2}{n\sum x_i^2}$$

حيث $\sum c_i = 0$ ؛ و $\sum c_i^2 = 1/\sum x_i^2$.

$$\operatorname{Var}\hat{b}_0 = \sigma_u^2\frac{\sum x_i^2 + n\bar{X}^2}{n\sum x_i^2} = \sigma_u^2\frac{\sum X_i^2 - n\bar{X}^2 + n\bar{X}^2}{n\sum x_i^2} = \sigma_u^2\frac{\sum X_i^2}{n\sum x_i^2}$$

حيث أننا رأينا فى (أ) أن $\sum x_i^2 = \sum X_i^2 - n\bar{X}^2$.

٦ - ١٥ بالنسبة لمشاهدات الاستهلاك - الدخل الإجماليين فى جدول ٦ - ٤ ، أوجد

(أ) S^2 ، (ب) $s_{\hat{b}_0}^2$ و $s_{\hat{b}_0}$ ، (جـ) $s_{\hat{b}_1}^2$ و $s_{\hat{b}_1}$.

(أ) الحسابات المطلوبة لإيجاد S^2 موضحة فى جدول ٦ - ٧ ، الذى يعتبر امتداداً لجدول ٦ - ٦ . نحصل على قيم $\hat{Y}_i$ الموجودة فى جدول ٦ - ٧ باحلال قيم X_i فى معادلة الانحدار السابق إيجادها فى المسألة ٦ - ٩ (أ) .

$$s^2 = \hat{\sigma}_u^2 = \frac{\sum e_i^2}{n-k} = \frac{115.2752}{12-2} = 11.52752 \cong 11.53$$

(ب)

$$s_{\hat{b}_0}^2 = \frac{\sum e_i^2}{n-k}\frac{\sum X_i^2}{n\sum x_i^2} = \frac{s\sum X_i^2}{n\sum x_i^2} \cong \frac{(11.53)(257,112)}{(12)(4,812)} \cong 51.34$$

فتكون :

$$s_{\hat{b}_0} = \sqrt{s_{\hat{b}_0}^2} \cong \sqrt{51.34} \cong 7.17$$

(جـ)

$$s_{\hat{b}_1}^2 = \frac{\sum e_i^2}{(n-k)\sum x_i^2} = \frac{s^2}{\sum x_i^2} \cong \frac{11.53}{4,812} \cong 0.0024$$

جدول ٦ – ٧ انحدار الاستهلاك : حسابات اختبار معنوية المعالم

n	Y_i	X_i	$\hat{Y}_i$	e_i	e_i^2	X_i^2	x_i^2
1	102	114	100.34	1.66	2.7556	12,996	961
2	106	118	103.78	2.22	4.9284	13,924	729
3	108	126	110.66	– 2.66	7.0756	15,876	361
4	110	130	114.10	– 4.10	16.8100	16,900	225
5	122	136	119.26	2.74	7.5076	18,496	81
6	124	140	122.70	1.30	1.6900	19,600	25
7	128	148	129.58	– 1.58	2.4964	21,904	9
8	130	156	136.46	– 6.46	41.7316	24,336	121
9	142	160	139.90	2.10	4.4100	25,600	225
10	148	164	143.34	4.66	21.7156	26,896	361
11	150	170	148.50	1.50	2.2500	28,900	625
12	154	178	155.38	– 1.38	1.9044	31,684	1,089
				$\sum e_i = 0$	$\sum e_i^2 = 115.2752$	$\sum X_i^2 = 257,112$	$\sum x_i^2 = 4,812$

فتكون $s_{\hat{b}_1} = \sqrt{s^2_{\hat{b}_1}} \cong \sqrt{0.0024} \cong 0.05$

٦ – ١٦ (أ) أذكر الفرض العدمى والفرض البديل لاختبار المعنوية الإحصائية لمعالم معادلة الانحدار المقدرة فى المسألة ٦ – ٩ (أ) . (ب) ما شكل توزيع المعاينة لكل من $\hat{b}_0$ و $\hat{b}_1$ ؟ (جـ) أى توزيع يجب استخدامه لاختبار المعنوية الإحصائية لكل من b_0 و b_1 ؟ (د) ما هى درجات الحرية ؟

(أ) لاختبار المعنوية الإحصائية لكل من b_0 و b_1 ، نستخدم الفرض العدمى H_0 ، والفرض البديل H_1 التاليين (أنظر قسم ٥ – ٢) :

$$H_0 : b_0 = 0 \quad : \quad \text{مقابل} \quad : \quad H_0 : b_0 \neq 0$$

$$H_0 : b_1 = 0 \quad : \quad \text{مقابل} \quad : \quad H_1 : b_1 \neq 0$$

المرجو من تحليل الانحدار أن نرفض الفرض H_0 وأن نقبل الفرض H_1 ، بأن كلا من b_0 و b_1 لا تساوى الصفر مستخدمين اختباراً له ذيلان .

(ب) حيث يفترض أن u_i تتبع التوزيع الطبيعى (فرض ١ فى قسم ٦ – ١) ، فإن Y_i أيضاً تتبع التوزيع الطبيعى (حيث أن X_i ثابتة فرضاً – الفرض 0) . وكنتيجة ، فإن $\hat{b}_0$ و $\hat{b}_1$ تكون أيضاً موزعة طبيعياً .

(جـ) لاختبار المعنوية الإحصائية لكل من b_0 و b_1 ، فإنه يجب استخدام توزيع t (من ملحق 0) لأن $\hat{b}_0$ و $\hat{b}_1$ تتبع التوزيع الطبيعى ، ولكن تباين b_0 وتباين b_1 غير معلومين (حيث أن σ_u^2 غير معلوم) وكذلك $n < 30$ (أنظر قسم ٤ – ٤) .

(د) درجات الحرية هى $n - k$ ، حيث n هى عدد المشاهدات و k هى عدد المعالم المقدرة . حيث أنه فى تحليل الانحدار البسيط يتم تقدير اثنين من المعالم ($\hat{b}_0$ و $\hat{b}_1$) ، تكون درجات الحرية $df = n - k = n - 2$. وتكون b_1 بين 0.75 ، 0.97 (أى أن $0.75 < b_1 < 0.97$) بدرجة ثقة 95% .

٦ – ١٧ اختبر عند مستوى معنوية (أ) b_0 ، (ب) b_1 فى المسألة ٦ – ٩ (أ) .

(أ)

$$t_0 = \frac{\hat{b}_0 - b_0}{s_{\hat{b}_0}} \cong \frac{2.30 - 0}{7.17} \cong 0.32$$

وحيث أن t_0 أصغر من القيمة الجدولية $t = 2.228$ عند مستوى 5% (اختبار له ذيلان) بدرجات حرية 10 . (من ملحق) . وننتهى إلى أن b_0 ليست معنوية إحصائياً عند مستوى معنوية 5% (بمعنى أنه لا يمكننا رفض الفرض H_0 ، بأن $b_0 = 0$) .

(ب)
$$t_1 = \frac{\hat{b}_1 - b_1}{s_{\hat{b}_1}} \cong \frac{0.86 - 0}{0.05} \cong 17.2$$

وبالتالى فإن b_1 معنوية إحصائياً عند مستوى معنوية 5% (وكذلك مستوى 1%) (بمعنى أنه لا يمكننا رفض الفرض H_1 بأن $b_1 \neq 0$) .

٦ – ١٨ كون فترة الثقة 95% لكل من (أ) b_0 ، (ب) b_1 فى تمرين ٦ – ٩ (أ) .

(أ) فترة الثقة 5% للمعلمة b_0 هى (أنظر قسم ٤ – ٤)

$$b_0 = \hat{b}_0 \pm 2.228 s_{\hat{b}_0} = 2.30 \pm 2.228(7.17) = 2.30 \pm 15.97$$

أى أن b_0 تقع بين 13.67 ، 18.27 بدرجة ثقة 95% . لاحظ اتساع فترة الثقة 95% للمعلمة b_0 (وغير ذات معنى) ، مما يعكس حقيقة أن $\hat{b}_0$ غير معنوية بدرجة كبيرة .

(ب) فترة الثقة 65% للمعلمة b_1 هى :

$$b_1 = \hat{b}_1 \pm 2.228 s_{\hat{b}_1} = 0.86 \pm 2.228(0.05) = 0.86 \pm 0.11$$

أى أن b_1 تقع بين 0.75 ، 0.97 (أى $0.75 < b_1 < 0.97$) بدرجة ثقة 95% .

اختبار جودة التوفيق والارتباط :

٦ – ١٩ اشتق صيغة فى R^2

يعرف معامل التحديد ، R^2 بأنه نسبة التغير الإجمالى فى Y « الذى يفسره » انحدار Y على X . أما التغير الإجمالى فى Y فهو إجمالى مجموع المربعات $TSS = \sum(\hat{Y}_i - \bar{Y})^2 | = \sum \hat{y}_i^2$. التغير المفسر فى Y أو مجموع مربعات الانحدار $RSS = \sum(Y_i - \bar{Y})^2 = \sum y_i^2$. بتغير البواقى فى Y أو مجموع مربعات الخطأ $ESS = \sum(Y_i - \hat{Y}_i)^2 = \sum e_i^2$.

$$\begin{aligned}
TSS &= RSS + ESS \\
\sum(Y_i - \bar{Y})^2 &= \sum(\hat{Y}_i - \bar{Y})^2 + \sum(Y_i - \hat{Y}_i)^2 \\
\sum y_i^2 &= \sum \hat{y}_i^2 + \sum e_i^2
\end{aligned}$$

بقسمة الطرفين على $\sum y_i^2$ ، نحصل على :

$$\frac{\sum y_i^2}{\sum y_i^2} = \frac{\sum \hat{y}_i^2}{\sum y_i^2} + \frac{\sum e_i^2}{\sum y_i^2}$$

$$1 = \frac{\sum \hat{y}_i^2}{\sum y_i^2} + \frac{\sum e_i^2}{\sum y_i^2}$$

وبالتالى :

$$R^2 = \frac{\sum \hat{y}_i^2}{\sum y_i^2} = 1 - \frac{\sum e_i^2}{\sum y_i^2}$$

R^2 ليس لها تمييز وتقع $0 < R^2 \leqslant 1$ لأن $0 \leqslant ESS \leqslant TSS$. $R^2 = 0$ عندما ، على سبيل المثال، تقع كل نقاط العينة على خط مستقيم $Y = \bar{Y}$ أو على دائرة . $R^2 = 1$ عندما تقع كل نقاط العينة على خط الانحدار المقدر ، مشيراً إلى علاقة تامة .

٦ — ٢٠ ماذا يقيس معامل الارتباط ؟ ما هو مدى قيم هذا المعامل ؟ (ب) ما هى العلاقة بين الارتباط وتحليل الانحدار ؟

(أ) يقيس معامل الارتباط درجة الاقتران بين متغيرين أو أكثر . فى حالة المتغيرين ، فان معامل الارتباط الخطى البسيط ، r لمجموعة مشاهدات عينة هو

$$r = \sqrt{R^2} = \frac{\sum x_i y_i}{\sqrt{\sum x_i^2}\sqrt{\sum y_i^2}} = \sqrt{\hat{b}_1 \frac{\sum x_i y_i}{\sum y_i^2}}$$

$-1 < r < +1$. $r < 0$ تعنى أن X و Y يتحركان فى اتجاهات عكسية مثل ، على سبيل المثال ، الكمية المطلوبة من سلعة ما وسعرها . $r > 0$ تشير إلى أن X و Y يتغيران فى نفس الاتجاه ، مثل الكمية المعروضة من سلعة ما وسعرها. $r = -1$ تشير إلى ارتباط عكسى تام (بمعنى أن كل مشاهدات العينة تقع على خط مستقيم ذى ميل سالب) ؛ ولكن $r = 1$ تشير إلى ارتباط تام موجب (أى أن كل مشاهدات العينة تقع على خط مستقيم ذى ميل موجب) . ونادرا ما تكون $r = \pm 1$. وكلما قربت r من ± 1 ، كلما زادت درجة العلاقة الخطية الموجبة أو السالبة . ويجب ملاحظة أن إشارة r هى نفس إشارة $\hat{b}_1$. ويعنى معامل ارتباط صفر أنه لا توجد علاقة خطية من أى نوع بين X و Y (بمعنى أنهما يميلان إلى التحرك بدون صلة بينهما . فمثلا ، إذا كانت مشاهدات العينة تقع تماماً على دائرة ، فإن هناك علاقة تامة غير خطية ولكن علاقة خطية صفرية وتكون $r = 0$.

(ب) ينطوى تحليل الانحدار (ولكنه لا يثبت) على وجود علاقة سببية بين المتغير المستقل ، X ، والمتغير التابع ، Y . تحليل الارتباط لا ينطوى على أى سببية أو تبعية ولكنه ببساطة يشير إلى نوع ودرجة الاقتران بين متغيرين . فعلا ولكن ، يمكن أن يكون بين X و Y ارتباط مرتفع بسبب أن متغيرا ثالث يؤثر بشدة على كليهما معا . لهذا فإن تحليل الارتباط هو أداة أقل قوة من تحليل الانحدار . ونادراً ما يستخدم بمفرده فى الحياة العملية . فى الواقع ، فإن الاستخدام الرئيسى لتحديد الارتباط هو لتحديد درجة الاقتران الموجودة فى تحليل الانحدار . وهذا يعطى بمعامل التحديد ، وهو مربع معامل الارتباط .

٦ — ٢١ اشتق المعادلة (أ) $r = \sum x_i y_i / (\sqrt{\sum x_i^2}\sqrt{\sum y_i^2})$ (ب) $r = \sqrt{\hat{b}_1(\sum x_i y_i / \sum y_i^2)}$.

ارشاد : ابدأ بتوضيح أن $\sum x_i y_i$ هى مقياس للاقتران بين X و Y (ب)

ارشاد : ابدأ بالمعادلة $r = \sum x_i y_i / (\sqrt{\sum x_i^2}\sqrt{\sum y_i^2})$.

(أ) تمدنا $\sum x_i y_i$ بمقياس للاقتران بين X ، Y لأنه إذا كانت X ، Y تزيدان أو تنقصان معاً فإن $x_i y_i > 0$ ، بينما إذا زادت X ونقصت Y أو العكس ، فإن $x_i y_i < 0$ ∴ إذا تضمنت كل أو معظم مشاهدات العينة زيادة أو نقصان فى X ، Y كليهما ، تكون $\sum x_i y_i$ وكبيرة ، مشيرة إلى ارتباط طردى كبير . إذا كانت كل أو معظم مشاهدات العينة تتضمن تغيرات عكسية فى X و Y تكون $\sum x_i y_i < 0$ وكبيرة ، مشيرة إلى ارتباط عكسى كبير . ولكن ، إذا كان بعض مشاهدات X و Y تتحرك فى نفس الاتجاه ، بينما البعض الآخر يتحرك فى الاتجاه المضاد ، فإن $\sum x_i y_i$ سوف تكون صغيرة ، مشيرة إلى صافى ارتباط موجب أو سالب . ولكن ، قياس درجة الاقتران باستخدام $\sum x_i y_i$ يكون له عيبان . الأول ، أنه كلما كبر عدد مشاهدات العينة ، كلما كبرت $\sum x_i y_i$. والثانى ، أن $\sum x_i y_i$ تكون معبرا عنها بوحدات المشكلة . ويمكن التغلب على هذه المشاكل بقسمة $\sum x_i y_i$ على n (عدد المشاهدات فى العينة) وعلى الانحراف المعيارى للمتغيرات X و Y $(\sqrt{\sum x_i^2/n}$ و $\sqrt{\sum y_i^2/n})$ وتصبح

$$\text{تغاير } X \text{ و } Y = \frac{\sum x_i y_i}{n} \qquad (٦ - ٣٠)$$

$$r = \frac{\sum x_i y_i}{\sqrt{\sum x_i^2}\sqrt{\sum y_i^2}} = \frac{\sum x_i y_i}{n\sqrt{\frac{\sum x_i^2}{n}}\sqrt{\frac{\sum y_i^2}{n}}}$$

(ب)
$$r = \frac{\sum x_i y_i}{\sqrt{\sum x_i^2}\sqrt{\sum y_i^2}} = \frac{\sqrt{\sum x_i y_i}}{\sqrt{\sum x_i^2}} \frac{\sqrt{\sum x_i y_i}}{\sqrt{\sum y_i^2}} = \sqrt{\hat{b}_1 \frac{\sum x_i y_i}{\sum y_i^2}}$$

٦ – ٢٢ أوجد R^2 لانحدار الاستهلاك المقدر فى المسألة ٦ – ٩ باستخدام المعادلة (أ) $R^2 = \sum \hat{y}_i^2 / \sum y_i^2$ (ب) $R^2 = 1 - \sum e_i^2 / \sum y_i^2$

(أ) نعرف من المسألة ٦ – ١٩ أن $\sum y_i^2 = \sum \hat{y}_i^2 + \sum e_i^2$ ، فتكون $\sum \hat{y}_i^2 = \sum y_i^2 - \sum e_i^2$ وحيث أن $\sum y_i^2 = 3{,}684$ (بتربيع وإضافة عناصر y_i's من جدول ٦ – ٦) ، $\sum e_i^2 = 115.2572$ (من جدول ٦ – ٧) $\sum \hat{y}_i^2 = 3{,}684 - 115.2572$ فتكون $= 3{,}568.7428$

$$R^2 = \frac{\sum \hat{y}_i^2}{\sum y_i^2} = \frac{3{,}568.7428}{3{,}684} \cong 0.9687, \text{ or } 96.87\%$$

(ب) باستخدام $\sum e_i^2 = 115.2572$ و $\sum y_i^2 = 3{,}684$ ، نحصل على

$$R^2 = 1 - \frac{\sum e_i^2}{\sum y_i^2} = 1 - \frac{115.2572}{3{,}684} \cong 0.9687, \text{ or } 96.87\%$$

كما فى (أ)

٦ – ٢٣ أوجد r لانحدار الاستهلاك المقدر فى المسألة ٦ – ٩ باستخدام (أ) $\sqrt{R^2}$ ، (ب) $r = \sum x_i y_i / (\sqrt{\sum x_i^2}\sqrt{\sum y_i^2})$ و (جـ) $r = \sqrt{\hat{b}_1 (\sum x_i y_i / \sum y_i^2)}$

(أ) $r = \sqrt{R^2} \cong \sqrt{0.9687} \cong 0.9842$ وهى موجبة لأن $\hat{b}_1 > 0$.

(ب) باستخدام $\sum x_i y_i = 4{,}144$ ، $\sum x_i^2 = 4{,}812$ من جدول ٦ – ٦ ، $\sum y_i^2 = 3{,}684$ من تمرين ٦ – ٢٢ (أ) ، نحصل على :

$$r = \frac{\sum x_i y_i}{\sqrt{\sum x_i^2}\sqrt{\sum y_i^2}} \cong \frac{4{,}144}{\sqrt{4{,}812}\sqrt{3{,}684}} \cong 0.9841$$

والفرق الصغير جدا بين قيمة r هنا والقيمة السابق إيجادها فى (أ) يرجع إلى أخطاء التقريب .

(جـ) باستخدام $\hat{b}_1 \cong 0.86$ السابق إيجادها فى المسألة ٦ – ٩ (أ) ،

$$r = \sqrt{\hat{b}_1 \frac{\sum x_i y_i}{\sum y_i^2}} \cong \sqrt{\frac{(0.86)(4{,}144)}{3{,}684}} \cong 0.9836$$

٦ – ٢٤ (أ) أوجد معامل ارتباط الرتب أى معامل ارتباط سيرمان بين درجات أعمال السنة وترتيب اختبار الذكاء IQ لعينة عشوائية من 10 طلاب من فصل كبير ، كما هي موضحة بجدول ٦ – ٨ ، باستخدام معادلة (٦ – ٣١) . (ب) متى يستخدم ارتباط الرتب ؟

جدول ٦ – ٨ درجات أعمال السنة والترتيب فى اختبار IQ

الطالب	1	2	3	4	5	6	7	8	9	10
درجات أعمال السنة	77	78	65	84	84	88	67	92	68	96
ترتيب	7	6	8	5	4	3	9	1	10	2

$$r' = 1 - \frac{6\sum D^2}{n(n^2-1)} \qquad (٦ - ٣١) \quad (أ)$$

حيث D = الفرق فى رتبة العنصرين المتناظرين فى كل زوج (إما ترتيباً تنازلياً أو تصاعدياً ، مع تخصيص الرتبة المتوسطة للمشاهدات من ذات القيمة الواحدة)

n = عدد المشاهدات

ويتضمن جدول ٦ – ٩ الحسابات اللازمة لإيجاد r .

$$r' = 1 - \frac{6\sum D^2}{n(n^2-1)} = 1 - \frac{6(10.50)}{10(99)} = 1 - \frac{63}{990} \cong 0.94$$

جدول ٦ – ٩ الحسابات اللازمة لإيجاد معامل ارتباط الرتب

n	درجات أعمال السنة	رتبة أعمال السنة	رتبة IQ	D	D^2
1	96	1	2	−1	1
2	92	2	1	1	1
3	88	3	3	0	0
4	84	4.5	4	0.5	0.25
5	84	4.5	5	−0.5	0.25
[illegible]	78	6	6	0	0
	77	7	7	0	0
	68	8	10	−2	4
	67	9	9	0	0
10	65	10	8	2	4
					$\sum D^2 = 10.50$

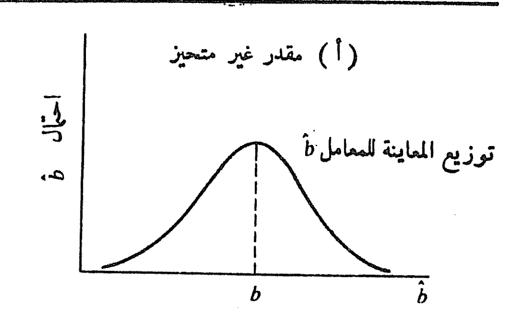
(ب) يستخدم ارتباط الرتب للبيانات الكيفية مثل المهنة ، التعليم ، الجنس ، الخ . عندما لا يمكن إيجاد معامل الارتباط لغياب القيم الرقية . كما يستخدم ارتباط الرتب عندما لا يكون متاحاً القيم الدقيقة لبعض أو كل المتغيرات (وبالتالى ، مرة أخرى ، لا يمكن إيجاد معامل الارتباط) . فضلا على ذلك ، فى حالة عدد كبير من المشاهدات ذات القيم العالية ، فيمكن إيجاد r' كتقدير للمعامل r لتجنب استخدام وقت طويل فى الحسابات (ولكن ، سهولة استخدام الكمبيوتر قد حذفت من الناحية العملية هذا السبب لاستخدام r') .

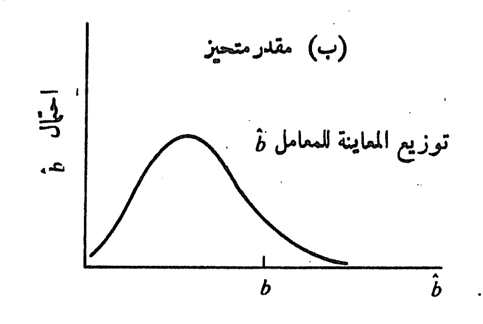
خصائص مقدرات المربعات الصغرى العادية :

٦ – ٢٥ (أ) ماذا يقصد بمقدر غير متحيز ؟ كيف يعرف التحيز ؟ (ب) ارسم شكلا يوضح توزيع المعاينة لمقدر غير متحيز وآخر متحيز .

(أ) يعتبر المقدر غير متحيز إذا كان وسط توزيع المعاينة الخاص به يساوى المعلمة الحقيقية . وسط توزيع المعاينة هو القيمة المتوقعة للمقدر . وغياب التحيز يعنى أن $E(\hat{b}) = b$ ، حيث $\hat{b}$ هى المقدر للمعلمة الحقيقية ، b . وعليه فيعرف التحيز بالفرق بين القيمة المتوقعة للمقدر وبين المعلمة الحقيقية . أى أن التحيز = $E(\hat{b}) - b$. لاحظ أن عدم وجود التحيز لا يعنى أن $\hat{b} = b$ ، ولكن فى المعاينة العشوائية المتكررة ، فإننا نحصل ، فى المتوسط ، على التقدير الصحيح . ونأمل أن العينة التى حصلنا عليها فعلا قريبة من وسط توزيع المعاينة المقدر .

(ب) يوضح شكل ٦ – ٧ (أ) توزيع المعاينة لمقدر غير متحيز ، كما يوضح شكل ٦ – ٧ (ب) توزيع المعاينة لمقدر متحيز .





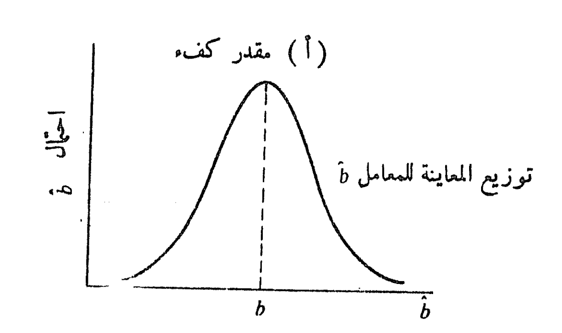
شكل ٦ – ٧

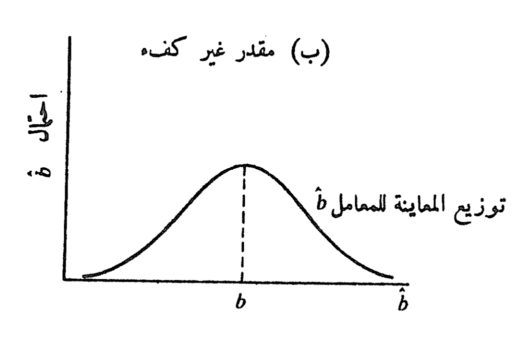
٦ – ٢٦ (أ) ماذا يقصد بأفضل مقدر غير متحيز أو مقدر كفء ؟ لماذا يكون هذا مهما ؟ (ب) ارسم شكلا لتوزيع المعاينة لمقدرين غير متحيزين ، أحدهما كفؤ .

(أ) أفضل مقدر غير متحيز أى مقدر كفؤ يشير إلى المقدر صاحب أصغر تباين بين المقدرات غير المتحيزة . فهو مقدر غير متحيز ذو توزيع أكثر تقارباً وأقل انتشارا من غيره .

وهذا مهم جداً لأن الباحث يكون أكثر تأكداً بأن المقدر أقرب إلى المعلمة الحقيقية للمجتمع موضع التقدير . أو بطريقة أخرى ، فإن المقدر الكفء يكون له أصغر فترة ثقة ومن المرجح أن يكون معنوياً إحصائياً عن غيره من المقدرات . ويجب أن يلاحظ أن أصغر تباين ليس مهما فى حد ذاته – إلا إذا اقترن بغياب التحيز .

(ب) يوضح شكل ٦ – ٨ (أ) توزيع المعاينة كمقدر كفء . بينما يوضح شكل ٦ – ٨ (ب) مقدراً غير كفء .





شكل ٦ – ٨

٦ – ٢٧ لماذا يكثر استخدام مقدرات OLS ؟ هل هى تمتاز عن غيرها من المقدرات ؟

يكثر استخدام مقدرات OLS لأنها BLUE (أفضل مقدرات خطية غير متحيزة) . أى ، من بين جميع المقدرات الخطية غير المتحيزة ، فإن لها أصغر تباين . وعادة ما يشار إلى خصائص BLUE هذه لمقدرات OLS كنظرية جاوس – ماركوف » . ولكن ، قد تمتاز مقدرات غير خطية عن OLS (بمعنى أنها يمكن أن تكون غير متحيزة ولها تباين أقل) . وحيث أنه عادة من الصعب أو من المستحيل إيجاد تباين أو تحيز المقدرات غير الخطية ، تبقى مقدرات OLS الأكثر شيوعاً فى الاستخدام . ومقدرات OLS ، بوصفها خطية ، تكون أسهل فى الاستخدام من المقدرات غير الخطية .

٦ – ٢٨ ماذا يقصد بمتوسط مربع الخطأ ؟ لماذا ومتى يكون استخدام قاعدة « النهاية الصغرى لمتوسط مربع الخطأ » مفيداً ؟ (ب) أثبت أن متوسط مربع الخطأ يساوى التباين زائداً مربع تحيز المقدر .

(أ) $$\text{MSE}(\hat{b}) = E(\hat{b} - b)^2 = \text{var}\,\hat{b} + (\text{bias}\,\hat{b})^2$$

وتنشأ قاعدة القيمة الصغرى للمقدر MSE عندما يواجه الباحث مقدراً متحيزا قليلا ولكن تباينه أصغر من أى مقدر آخر غير متحيز . فعلى الأرجح ، سوف يختار الباحث المقدر صاحب أصغر MSE . هذه القاعدة تأخذ من التباين الكبير ومربع التحيز الكبير موقفاً واحداً . ولكن ، يستخدم هذا فقط عندما يكون لمقدر OLS تباين « كبير بدرجة غير مقبولة » .

(ب)

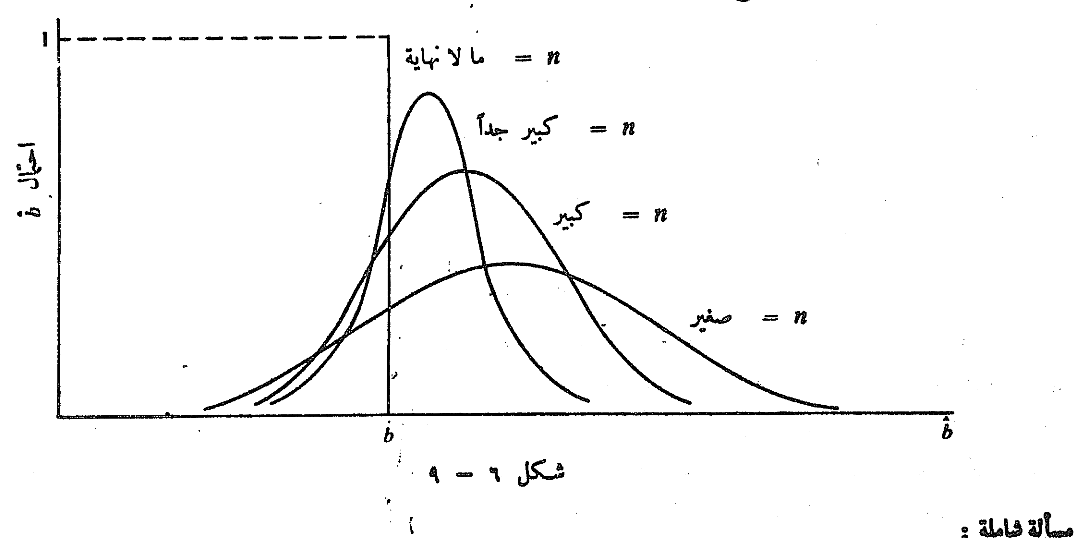
$$\begin{aligned}\text{MSE}(\hat{b}) &= E(\hat{b}-b)^2\\ &= E[\hat{b}-E(\hat{b})+E(\hat{b})-b]^2\\ &= E[\hat{b}-E(\hat{b})]^2+[E(\hat{b})-b]^2+2E\{[\hat{b}-E(\hat{b})][E(\hat{b})-b]\}\\ &= \text{var}\,\hat{b}+(\text{bias}\,\hat{b})^2\end{aligned}$$

لأن $E[\hat{b}-E(\hat{b})]^2=\text{var}\,\hat{b},\ [E(\hat{b})-b]^2=(\text{bias}\,\hat{b})^2,$ و $E\{[\hat{b}-E(\hat{b})][E(\hat{b})-b]\}=0$ لأن
لأنها تساوى $E\{\hat{b}E(\hat{b})-[E(\hat{b})]^2-\hat{b}b+bE(\hat{b})\}=[E(\hat{b})]^2-[E(\hat{b})]^2-bE(\hat{b})+bE(\hat{b})$
$=0.$

٦ – ٢٩ (أ) ماذا يقصد بالاتساق ؟ (ب) ارسم شكلا لتوزيع المعاينة لمقدر متسق .

(أ) هناك شرطان لكى يكون المقدر متسقاً : (١) مع كبر حجم العينة ، فإن المقدر يجب أن يقترب أكثر فأكثر من المعلمة الحقيقية (ويشار إلى هذا كعدم تحيز فى اللانهاية) (٢) مع اقتراب حجم العينة من ما لا كنهاية نهاية ، فإن توزيع المعاينة للمقدار يجب أن ينتهى أو يصبح خطاً مستقيما رأسياً بارتفاع (احتمال) ١ [illegible] القيمة الحقيقية للمعلمة . وخاصية الاتساق هذه للعينات الكبيرة تستخدم فقط فى حالات عندما لا يمكن [illegible] العينات الصغيرة أو المقدرات ذات أصغر MSE .

(ب) فى شكل ٦ – ٩ ، $\hat{b}$ مقدر متسق للمعلمة b لأنه مع تزايد n تقترب $\hat{b}$ من b ، ومع اقتراب n من ما لا نهاية كنهاية ، فإن توزيع المعاينة للمعامل $\hat{b}$ ينتهى إلى b .



شكل ٦ – ٩

مسألة شاملة :

٦ – ٣٠ يعطى جدول ٦ – ١٠ دخل الفرد الحقيقى ، لأقرب 1.000 دولار أمريكى ، Y_i فى 15 لمقدولة متقدمة والنسبة المناظرة لقوة العمل فى الزراعة ، X_i ، لأقرب 1% فى عام ١٩٨١ . (أ) قدر معادلة انحدار Yi على X_i (ب) اختبر عند مستوى معنوية 5% المعنوية الاحصائية للمعالم . (جـ) أوجد معامل التحديد (د) ضع نتائج (أ) فى صورة موجزة قياسية .

جدول ٦ — ١٠ دخل الفرد ، Y_i (بالألف دولار) ، ونسبة القوة العاملة فى الزراعة X_i ، فى 15 دولة متقدمة فى ١٩٨١ .

الدولة	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Y_i	6	8	8	7	7	12	9	8	9	10	10	11	9	10	11
X_i	9	10	8	7	10	4	5	5	6	8	7	4	9	5	8

(أ) تستخدم الأعمدة السبعة الأولى فى جدول ٦ — ١١ للاجابة على (أ) . ويتم ملء باقى الجدول باستخدام نتائج (أ) لإجابة (ب) ، (ج) من المسألة .

$$\hat{b}_1 = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{-28}{60} \cong -0.47$$

$$\hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X} = 9 - (-0.47)(7) \cong 12.29$$

$$\hat{Y}_i = 12.29 - 0.47 X_i$$

جدول ٦ — ١١ مسودة

الدولة	Y_i	X_i	y_i	x_i	$x_i y_i$	x_i^2	$\hat{Y}_i$	e_i	e_i^2	X_i^2	y_i^2
1	6	9	−3	2	−6	4	8.06	−2.06	4.2436	81	9
2	8	10	−1	3	−3	9	7.59	0.41	0.1681	100	1
3	8	8	−1	1	−1	1	8.53	−0.53	0.2809	64	1
4	7	7	−2	0	0	0	9.00	−2.00	4.0000	49	4
5	7	10	−2	3	−6	9	7.59	−0.59	0.3481	100	4
6	12	4	3	−3	−9	9	10.41	1.59	2.5281	16	9
7	9	5	0	−2	0	4	9.94	−0.94	0.8836	25	0
8	8	5	−1	−2	2	4	9.94	−1.94	3.7636	25	1
9	9	6	0	−1	0	1	9.47	−0.47	0.2209	36	0
10	10	8	1	1	1	1	8.53	1.47	2.1609	64	1
11	10	7	1	0	0	0	9.00	1.00	1.0000	49	1
12	11	4	2	−3	−6	9	10.41	0.59	0.3481	16	4
13	9	9	0	2	0	4	8.06	0.94	0.8836	81	0
14	10	5	1	−2	−2	4	9.94	0.06	0.0036	25	1
15	11	8	2	1	2	1	8.53	2.47	6.1009	64	4
$n = 15$	$\sum Y_i = 135$ $\bar{Y} = 9$	$\sum X_i = 105$ $\bar{X} = 7$	$\sum y_i = 0$	$\sum x_i = 0$	$\sum x_i y_i = -28$	$\sum x_i^2 = 60$		$\sum e_i = 0$	$\sum e_i^2 = 26.9340$	$\sum X_i^2 = 795$	$\sum y_i^2 = 40$

(ب)

$$s_{\hat{b}_0}^2 = \frac{\sum e_i^2}{n-k} \frac{\sum X_i^2}{n \sum x_i^2} = \frac{(26.9340)(795)}{(15-2)(15)(60)} \cong 1.83 \quad \text{and} \quad s_{\hat{b}_0} \cong 1.35$$

$$s_{\hat{b}_1}^2 = \frac{\sum e_i^2}{(n-k)\sum x_i^2} = \frac{26.9340}{(15-2)(60)} \cong 0.03 \quad \text{and} \quad s_{\hat{b}_1} \cong 0.17$$

$$t_0 = \frac{\hat{b}_0}{s_{\hat{b}_0}} \cong \frac{12.29}{1.35} \cong 9.10$$

$$t_1 = \frac{\hat{b}_1}{s_{\hat{b}_1}} \cong \frac{-0.47}{0.17} \cong -2.76$$

وبالتالى ، فإن كلا من $\hat{b}_0$ و $\hat{b}_1$ معنوية إحصائياً عند مستوى معنوية 5% .

(جـ)
$$R^2 = 1 - \frac{\sum e_i^2}{\sum y_i^2} = 1 - \frac{26.9340}{40} \cong 0.33$$

(د)
$$\hat{Y}_i = \underset{(9.10)}{12.29} - \underset{(-2.76)}{0.47X_i} \qquad . \; R^2 = 0.33$$

الأرقام داخل الأقواس تحت المعالم المقدرة تشير إلى قيم (t) المناظرة . كطريقة بديلة يمكن كتابة الخطأ المعيارى للتقدير داخل الأقواس .

مسائل إضافية

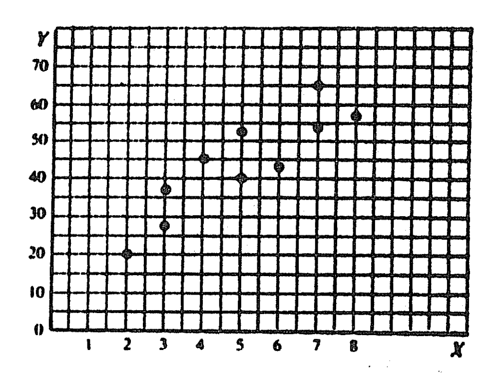
النموذج الخطى ذو المتغيرين :

٦ – ٣١ ارسم شكل انتشار لبيانات جدول ٦ – ١٢ وحدد بالنظر إذا كان هناك علاقة خطية تقريبية بين Y_i ، X_i .

الإجابة : العلاقة بين X ، Y فى شكل ٦ – ١٠ خطية تقريباً .

جدول ٦ – ١٢ مشاهدات من المتغيرات Y و X

n	Y_i	X_i
1	20	2
2	28	3
3	40	5
4	45	4
5	37	3
6	52	5
7	54	7
8	43	6
9	65	7
10	56	8



شكل ٦ – ١٠

٦ – ٣٢ اذكر فروض نموذج الانحدار الكلاسيكي (OLS) في صورة رياضية .

الإجابة :

$$u \sim N(0, \sigma_u^2) \quad (٦ - ٢١)$$

$$E(u_i u_j) = 0 \quad \text{for } i \neq j; \quad i, j = 1, 2, \ldots, n \quad (٦ - ٢٢)$$

$$E(X_i u_i) = 0 \quad (٦ - ٢٣)$$

(انظر المسألة ٦ – ٤) .

طريقة المربعات الصغرى العادية :

٦ – ٣٣ عبر رياضيا عن العبارات والصيغ الآتية : (أ) أوجد القيمة الصغرى لمجموع مربعات انحرافات كل قيمة Y عن القيمة التوفيقية المناظرة لها .

(ب) أوجد القيمة الصغرى لمجموع مربعات البواقى . (جـ) المعادلات الطبيعية (د) الصيغ لتقدير $\hat{b}_1$ و $\hat{b}_0$.

الإجابة : (أ) $\text{Min} \sum (Y_i - \hat{Y}_i)^2$ (ب) $\sum e_i^2$ (جـ) $\sum Y_i = n\hat{b}_0 + \hat{b}_1 \sum X_i$ و $\sum X_i Y_i = \hat{b}_0 \sum X_i + \hat{b}_1 \sum X_i^2$

(د) $\hat{b}_1 = (n\sum X_i Y_i - \sum X_i \sum Y_i)/[n\sum X_i^2 - (\sum X_i)^2] = \sum x_i y_i / \sum x_i^2$ و $\hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X}$

٦ – ٣٤ بالنسبة لبيانات جدول ٦ – ١٢ ، أوجد قيمة (أ) b_1 ، (ب) b_0 .

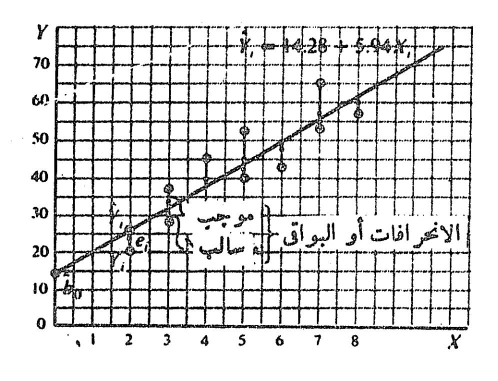
(جـ) اكتب معادلة خط انحدار OLS المقدر . الإجابة (أ) $\hat{b}_1 \cong 5.94$ (ب) $\hat{b}_0 \cong 14.28$ (جـ) $\hat{Y}_i = 14.28 + 5.94X_i$

٦ – ٣٥ (أ) ارسم على مجموعة من المحاور بيانات جدول ٦ – ١٢ ، خط انحدار OLS المقدر فى مسألة ٦ – ٣٤ ، ووضح البواقى .

(ب) وضح بالرسم أن خط الانحدار يمر بالنقطة $\bar{X}\bar{Y}$.

الإجابة : (أ) انظر شكل ٦ – ١١ (ب) عند $X_i = 5 = \bar{X}$ ، $\hat{Y}_i = 14.28 + 5.94(5) = 43.98 \cong \bar{Y} = 44$

(الفرق البسيط ناتج عن التقريب)



شكل ٦ – ١١

٦ – ٣٦ بالإشارة إلى خط انحدار OLS المقدر فى المسألة ٦ – ٣٤ ، اذكر (أ) معنى $\hat{b}_1$ ، (ب) معنى $\hat{b}_0$ ، (جـ) مرونة Y بالنسبة إلى X عند المتوسطات .

الإجابة : (أ) $\hat{b}_0$ هى (الجزء المقطوع من Y) (ب) $\hat{b}_1$ هى ميل خط انحدار OLS المقدر (جـ) $\eta \cong 0.68$

اختبارات معنوية تقديرات المعالم :

٦ – ٣٧ بالنسبة لبيانات جدول ٦ – ١٢ فى مسألة ٦ – ٣١ ، أوجد (أ) S^2 (ب) $s_{\hat{b}_0}^2$ و $s_{\hat{b}_0}$ ، (جـ) $s_{\hat{b}_1}^2$ و $s_{\hat{b}_1}$

الإجابة : (أ) $s^2 \cong 46.97$ (ب) $s_{\hat{b}_0}^2 \cong 37.31$ و $s_{\hat{b}_0} \cong 6.11$ (جـ) $s_{\hat{b}_1}^2 \cong 1.31$ و $s_{\hat{b}_1} \cong 1.14$

٦ – ٣٨ اختبر عند مستوى معنوية 5% كلا من (أ) b_0 ، (ب) b_1 فى المسألة ٦ – ٣٤ .

الإجابة : (أ) b_0 معنوية إحصائياً عند مستوى 5% (ب) b_1 أيضاً معنوية إحصائياً عند مستوى 5% .

٦ – ٣٩ كون فترة الثقة 95% لكل من (أ) b_0 ، (ب) b_1 فى مسألة ٦ – ٣٤ .

الإجابة : (أ) $0.19 < b_0 < 28.37$ (ب) $3.31 < b_1 < 8.57$

اختبار جودة التوفيق والارتباط :

٦ – ٤٠ بالنسبة لبيانات معادلة انحدار OLS المقدرة فى مسألة ٦ – ٣٤ ، أوجد (أ) R^2 (ب) r .

الإجابة : (أ) $R^2 \cong 0.77$ (ب) $r \cong 0.88$

٦ – ٤١ أوجد معامل ارتباط الرتب لعينة مشاهدات XY فى جدول ٦ – ١٢ .

الإجابة : $r' \cong 0.90$ $(\cong r \cong 0.88)$

خصائص مقدرات المربعات الصغرى العادية :

٦ – ٤٢ بالإشارة إلى $\hat{b}_0$ و $\hat{b}_1$ فى المسألة ٦ – ٣٤ ، هل هى : (أ) BLUE ؟ (ب) غير متحيزة فى اللانهاية ؟ (جـ) متسقة ؟

الإجابة : (أ) نعم (ب) نعم (جـ) نعم

٦ – ٤٣ بالإشارة إلى $\hat{b}_0$ و $\hat{b}_1$ فى المسألة ٦ – ٣٤ : (أ) ما هو MSE ؟ (ب) هل تعطى $\hat{b}_0$ و $\hat{b}_1$ أصغر قيمة للمقدار MSE ؟

الإجابة : (أ) $\text{MSE}(\hat{b}_0) = \text{var}\,\hat{b}_0$ و $\text{MSE}(\hat{b}_1) = \text{var}\,\hat{b}_1$ (ب) نعم

مسألة شاملة :

٦ – ٤٤ يعطى جدول ٦ – ١٣ بيانات عينة عشوائية من 12 عائلة من عدد الأطفال فى الأسرة Y_i ، وعدد الأطفال الذين قالوا وقت الزواج إنهم يرغبون فى إنجابه X_i . أوجد انحدار Y_i على X_i وضع النتائج فى صورة موجزة .

جدول ٦ – ١٣ عدد الأطفال فى الأسرة والعدد المرغوب من الأطفال

العائلة	1	2	3	4	5	6	7	8	9	10	11	12
Y_i	4	3	0	4	4	3	0	4	3	1	3	1
X_i	3	3	0	2	2	3	0	3	2	1	3	2

الإجابة :

$$\hat{Y}_i = \underset{(0.39)}{0.22} + \underset{(4.56)}{1.14}X_i \qquad R^2 = 0.68$$

والأرقام داخل الأقواس هى قيم t . وعليه فإن $\hat{b}_1$ معنوية إحصائياً عند مستوى معنوية 5% ، 1% ، ولكن $\hat{b}_0$ ليست معنوية .

الفصل السابع

تحليل الانحدار المتعدد

٧–١ النموذج الخطى لثلاثة متغيرات

يستخدم تحليل الانحدار المتعدد لاختبار الفروض عن العلاقة بين متغير تابع ، Y ، وإثنين أو أكثر من المتغيرات المستقلة ،.... $X_1, X_2, \ldots$ ، والتنبؤ . ويمكن كتابة نموذج الانحدار الثلاثى كالآتى :

$$Y_i = b_0 + b_1 X_{1i} + b_2 X_{2i} + u_i \qquad (٧ - ١)$$

الفرض الإضافى (إلى فروض النموذج الخطى البسيط) أنه لا توجد علاقة خطية تامة بين المتغيرات المستقلة $X_1, X_2, \ldots$.

ويمكن الحصول على تقديرات معالم المربعات الصغرى العادية OLS بإيجاد النهاية الصغرى لمجموع مربعات البواقى .

$$\sum e_i^2 = \sum (Y_i - \hat{Y}_i)^2 = \sum (Y_i - \hat{b}_0 - \hat{b}_1 X_{1i} - \hat{b}_2 X_{2i})^2$$

ويعطى هذا المعادلات الطبيعية الثلاث الآتية (انظر المسألة ٧ – ٢) :

$$\sum Y_i = n\hat{b}_0 + \hat{b}_1 \sum X_{1i} + \hat{b}_2 \sum X_{2i} \qquad (٧ - ٢)$$

$$\sum X_{1i} Y_i = \hat{b}_0 \sum X_{1i} + \hat{b}_1 \sum X_{1i}^2 + \hat{b}_2 \sum X_{1i} X_{2i} \qquad (٧ - ٣)$$

$$\sum X_{2i} Y_i = \hat{b}_0 \sum X_{2i} + \hat{b}_1 \sum X_{1i} X_{2i} + \hat{b}_2 \sum X_{2i}^2 \qquad (٧ - ٤)$$

والتى (عندما يعبر عنها فى صورة انحرافات المتغيرات عن متوسطاتها) يمكن حلها آنياً لإيجاد $\hat{b}_1$ و $\hat{b}_2$ معطية (انظر تمرين ٧ – ٣)

$$\hat{b}_1 = \frac{(\sum x_1 y)(\sum x_2^2) - (\sum x_2 y)(\sum x_1 x_2)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2} \qquad (٧ - ٥)$$

$$\hat{b}_2 = \frac{(\sum x_2 y)(\sum x_1^2) - (\sum x_1 y)(\sum x_1 x_2)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2} \qquad (٧ - ٦)$$

وتكون

$$\hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X}_1 - \hat{b}_2 \bar{X}_2 \qquad (٧ - ٧)$$

ويقيس المقدر $\hat{b}_1$ التغير فى Y بالنسبة لتغير مقداره الوحدة فى X_1 مع تثبيت X_2 . وتعرف $\hat{b}_2$ على نفس النمط . وتسمى المقدرات $\hat{b}_1$ و b_2 بمعاملات الانحدار الجزئية وتكون $\hat{b}_0$ ، $\hat{b}_1$ و $\hat{b}_2$ BLUE (انظر قسم ٦ – ٥) .

مثال ١ : جدول ٧ – ١ هو امتداد لجدول ٦ – ١ ويعطى عدد بوشلات الحنطة للأكر ، Y ، الناتج من استخدام كميات مختلفة من الأسمدة X_1 ، وكميات مختلفة من المبيدات الحشرية X_2 ، معبراً عنها بعدد الأرطال للأكر ، من عام ١٩٧١ إلى ١٩٨٠ . باستخدام معادلات (٧ – ٥) ، (٧ – ٦) ، (٧ – ٧) ، نحصل على

$$\hat{b}_1 = \frac{(\sum x_1 y)(\sum x_2^2) - (\sum x_2 y)(\sum x_1 x_2)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2} = \frac{(956)(504) - (900)(524)}{(576)(504) - (524)^2} \cong 0.65$$

$$\hat{b}_2 = \frac{(\sum x_2 y)(\sum x_1^2) - (\sum x_1 y)(\sum x_1 x_2)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2} = \frac{(900)(576) - (956)(524)}{(576)(504) - (524)^2} \cong 1.11$$

$$\hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X}_1 - \hat{b}_2 \bar{X}_2 \cong 57 - (0.65)(18) - (1.11)(12) \cong 31.98$$

جدول ٧ – ١ الحنطة المنتجة من السماد والمبيد المستخدم مع حسابات تقديرات المعالم

السنة	Y	X_1	X_2	y	x_1	x_2	x_1y	x_2y	x_1x_2	x_1^2	x_2^2
1971	40	6	4	−17	−12	−8	204	136	96	144	64
1972	44	10	4	−13	−8	−8	104	104	64	64	64
1973	46	12	5	−11	−6	−7	66	77	42	36	49
1974	48	14	7	−9	−4	−5	36	45	20	16	25
1975	52	16	9	−5	−2	−3	10	15	6	4	9
1976	58	18	12	1	0	0	0	0	0	0	0
1977	60	22	14	3	4	2	12	6	8	16	4
1978	68	24	20	11	6	8	66	88	48	36	64
1979	74	26	21	17	8	9	136	153	72	64	81
1980	80	32	24	23	14	12	322	276	168	196	144
$n = 10$	$\sum Y = 570$ $\bar{Y} = 57$	$\sum X_1 = 180$ $\bar{X}_1 = 18$	$\sum X_2 = 120$ $\bar{X}_2 = 12$	$\sum y = 0$	$\sum x_1 = 0$	$\sum x_2 = 0$	$\sum x_1 y = 956$	$\sum x_2 y = 900$	$\sum x_1 x_2 = 524$	$\sum x_1^2 = 576$	$\sum x_2^2 = 504$

وعليه فإن ، $\hat{Y}_i = 31.98 + 0.65\,X_{1i} + 1.10\,X_{2i}$. لتقدير معالم الانحدار لثلاثة متغيرات مستقلة أو أكثر ، انظر برنامج الكمبيوتر للمسألة ٧ – ٢٢ .

٧–٢ اختبارات معنوية لتقديرات المعالم

لاختبار المعنوية الاحصائية لتقديرات المعالم للانحدار المتعدد ، فإن تباين التقديرات مطلوب :

$$\mathrm{Var}\,\hat{b}_1 = \sigma_u^2 \frac{\sum x_2^2}{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2} \qquad (٧ - ٨)$$

$$\mathrm{Var}\,\hat{b}_2 = \sigma_u^2 \frac{\sum x_1^2}{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2} \qquad (٧ - ٩)$$

(عادة b_0 ليست موضع اهتمام أساسى ، أنظر المسألة ٧ – ٦ (هـ)) . وحيث أن σ_u^2 غير معلومة ، فإن تباين البواقى ، S^2 يستخدم كتقدير غير متحيز للتباين σ_u^2 :

$$s^2 = \hat{\sigma}_u^2 = \frac{\sum e_i^2}{n-k} \qquad (٦ - ١٢)$$

حيث K = عدد المعالم المقدرة

فتكون التقديرات غير المتحيزة لتباين $\hat{b}_0$ و $\hat{b}_1$ هى :

$$s_{\hat{b}_1}^2 = \frac{\sum e_i^2}{n-k} \frac{\sum x_2^2}{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2} \qquad (٧ - ١٠)$$

$$s_{\hat{b}_2}^2 = \frac{\sum e_i^2}{n-k} \frac{\sum x_1^2}{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2} \qquad (٧ - ١١)$$

وبالتالى فإن $S\hat{b}_1$ و $S\hat{b}_2$ هى الأخطاء المعيارية للتقديرات . وتجرى اختبارات الفروض عن b_1 و b_2 كما فى قسم ٦ – ٣ .

مثال ٢ : جدول ٧ – ٢ (وهو امتداد لجدول ٧ – ١) يبين الحسابات الإضافية اللازمة لاختبار المعنوية الإحصائية لكل من b_1 و b_2 . ويتم الحصول على قيم $\hat{Y}_i$ فى جدول ٧ – ٢ بالتعويض عن قيم X_{1i} و X_{2i} فى معادلة انحدار OLS المقدرة السابق إيجادها فى مثال ١ . (ويتم الحصول على قيم y_i^2 بتربيع y_i من جدول ٧ – ١ وسوف تستخدم فى قسم ٧ – ٢) .

جدول ٧ – ٢ حسابات الحنطة – السماد – المبيد لاختبار معنوية المعالم

السنة	Y	X_1	X_2	$\hat{Y}$	e	e^2	y^2
1971	40	6	4	40.32	− 0.32	0.1024	289
1972	44	10	4	42.92	1.08	1.1664	169
1973	46	12	5	45.33	0.67	0.4489	121
1974	48	14	7	48.85	− 0.85	0.7225	81
1975	52	16	9	52.37	− 0.37	0.1369	25
1976	58	18	12	57.00	1.00	1.0000	1
1977	60	22	14	61.82	− 1.82	3.3124	9
1978	68	24	20	69.78	− 1.78	3.1684	121
1979	74	26	21	72.19	1.81	3.2761	289
1980	80	32	24	79.42	0.58	0.3364	529
$n = 10$					$\sum e = 0$	$\sum e^2 = 13.6704$	$\sum y^2 = 1{,}634$

باستخدام قيم جدول ٧ - ٢ ، ٧ - ١ نحصل على :

$$s_{\hat{b}_1}^2 = \frac{\sum e_i^2}{n-k} \frac{\sum x_2^2}{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2} = \frac{13.6704}{10-3} \frac{504}{(576)(504) - (524)^2} \cong 0.06 \quad \text{and} \quad s_{\hat{b}_1} \cong 0.24$$

$$s_{\hat{b}_2}^2 = \frac{\sum e_i^2}{n-k} \frac{\sum x_1^2}{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2} = \frac{13.6704}{10-3} \frac{576}{(576)(504) - (524)^2} = 0.07 \quad \text{and} \quad s_{\hat{b}_2} \cong 0.27$$

وعليه فإن $t_1 = \hat{b}_1/s_{\hat{b}_1} \cong 0.65/0.24 \cong 2.70$, و $t_2 = \hat{b}_2/s_{\hat{b}_2} = 1.11/0.27 \cong 4.11.$ وحيث أن كلا من t_1 و t_2 تتجاوز $t = 2.365$ بدرجات حرية 7 عند مستوى معنوية 5%(من ملحق ٥) ، فإن كلا من b_1 و b_2 معنوية عند مستوى معنوية 5% .

٧ــ٣ معامل التحديد المتعدد

يعرف معامل التحديد المتعدد ، R^2 ، بأنه نسبة التغير الإجمالي فى Y الذى « يفسره » الانحدار المتعدد للمتغير Y على المتغيرين X_1 و X_2 ، (وكما هو موضح فى قسم ٦ - ٤) يمكن حسابه كالآتى (أنظر المسألة ٧ - ١٠) .

$$R^2 = \frac{\sum \hat{y}_i^2}{\sum y_i^2} = 1 - \frac{\sum e_i^2}{\sum y_i^2} = \frac{\hat{b}_1 \sum yx_1 + \hat{b}_2 \sum yx_2}{\sum y^2}$$

وحيث إن إضافة متغيرات مستقلة أو مفسرة أخرى يرفع على الأرجح $RSS = \sum \hat{y}_i^2$ لنفس قيمة $TSS = \sum y_i^2$ (انظر قسم ٦ - ٤) ، فإن R^2 تزيد . فإذا أخذنا فى الاعتبار نقص عدد درجات الحرية مع إضافة متغيرات مستقلة اضافية ، فإن R^2 المعدلة $\bar{R}^2$ أو يمكن حسابها كالآتى (انظر المسألة ٧ - ١٢) :

$$\bar{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-k} \qquad (١٢ - ٧)$$

حيث n = عدد المشاهدات

k = عدد المعالم المقدرة

مثال ٣ : يمكن إيجاد R^2 لمثال الحنطة - السماد - المبيد باستخدام جدول ٧ - ٢ :

$$R^2 = 1 - \frac{\sum e_i^2}{\sum y_i^2} = 1 - \frac{13.6704}{1,634} \cong 1 - 0.0084 = 0.9916, \text{ or } 99.16\%$$

قارن هذا مع قيمة R^2 وقدرها 97.10% فى حالة الانحدار البسيط ، عندما استخدم السماد كمتغير مستقل وحيد .

$$\bar{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-k} = 1 - (1 - 0.9916)\frac{10-1}{10-3} = 1 - 0.0084(1.2857) = 0.9892, \text{ or } 98.92\%$$

٧ــ٤ اختبار المعنوية الكلية للانحدار

يمكن اختبار المعنوية الإجمالية للانحدار باستخدام نسبة التباين « المفسر » إلى التباين غير « المفسر » ، ويتبع هذا توزيع F بدرجات حرية $k - 1$ و $n - k$ حيث n عدد المشاهدات ، k عدد المعالم المقدرة (انظر قسم ٥ - ٥)

$$F_{k-1,n-k} = \frac{\sum \hat{y}_i^2/(k-1)}{\sum e_i^2/(n-k)} = \frac{R^2/(k-1)}{(1-R^2)/(n-k)} \qquad (١٣ - ٧)$$

فإذا تجاوزت نسبة F المحسوبة قيمة F الجدولية عند مستوى المعنوية ودرجات الحرية المحددة (من ملحق ٧) يقبل الفرض بأن معالم الانحدار ليست جميعها مساوية للصفر وأن R^2 تختلف جوهريا عن الصفر .

مثال ٤ - لاختبار المعنوية الكلية للانحدار المقدر فى مثال ١ بمستوى 5% يمكن استخدام $R^2 = 0.9916$ (من مثال ٣) بحيث

$$F_{2,7} = \frac{0.9916/2}{(1 - 0.9916)/7} \cong 413.17$$

وحيث أن قيمة F المحسوبة تفوق القيمة الجدولية $F=4.74$ عند مستوى معنوية 5% وحيث $df = 2, 7$ (من ملحق ٧) ، نقبل الفرض بأن b_1 و b_2 لا تساوى الصفر معاً وأن R^2 تختلف معنوياً عن الصفر .

٧ــ٥ معاملات الارتباط الجزئى

يقيس معامل الارتباط الجزئى صافى الارتباط بين المتغير التابع ومتغير مستقل بعد حذف التأثير المشترك (أى مع تثبيت (المتغيرات المستقلة الأخرى فى النموذج . فمثلا $r_{YX_1 \cdot X_2}$ هو الارتباط الجزئى بين Y و X_1 بعد حذف تأثير X_2 من كل من Y و X_1 (أنظر المسألة ٧ – ١٩ (أ)) :

$$r_{YX_1 \cdot X_2} = \frac{r_{YX_1} - r_{YX_2} r_{X_1 X_2}}{\sqrt{1 - r^2_{X_1 X_2}}\sqrt{1 - r^2_{YX_2}}} \qquad (٧ - ١٤)$$

$$r_{YX_2 \cdot X_1} = \frac{r_{YX_2} - r_{YX_1} r_{X_1 X_2}}{\sqrt{1 - r^2_{X_1 X_2}}\sqrt{1 - r^2_{YX_1}}} \qquad (٧ - ١٥)$$

حيث r_{YX_1} معامل الارتباط البسيط بين Y و X_1 ، ويعرف r_{YX_2} و $r_{X_1 X_2}$ على نفس النمط . وتتراوح معاملات الارتباط الجزئية بين -1 ، $+1$ (كما هو الحال بالنسبة لمعاملات الارتباط البسيط) ، ويكون لها نفس إشارة معلمة المجتمع المناظرة ، وتستخدم لتحديد الأهمية النسبية للمتغيرات المفسرة المختلفة فى الانحدار المتعدد .

مثال ٥ : بالتعويض بقيم جدول ٧ – ١ ، ٧ – ٢ فى معادلة (٦ – ١٨) لمعامل الارتباط البسيط ، نحصل على

$$r_{YX_1} = \frac{\sum x_1 y}{\sqrt{\sum x_1^2}\sqrt{\sum y^2}} = \frac{956}{\sqrt{576}\sqrt{1,634}} \cong 0.9854$$

$$r_{YX_2} = \frac{\sum x_2 y}{\sqrt{\sum x_2^2}\sqrt{\sum y^2}} = \frac{900}{\sqrt{504}\sqrt{1,634}} \cong 0.9917$$

$$r_{X_1 X_2} = \frac{\sum x_2 x_1}{\sqrt{\sum x_2^2}\sqrt{\sum x_1^2}} = \frac{524}{\sqrt{504}\sqrt{576}} \cong 0.9725$$

وعليه

$$r_{YX_1 \cdot X_2} = \frac{r_{YX_1} - r_{YX_2} r_{X_1 X_2}}{\sqrt{1 - r^2_{X_1 X_2}}\sqrt{1 - r^2_{YX_2}}} = \frac{0.9854 - (0.9917)(0.9725)}{\sqrt{1 - 0.9725^2}\sqrt{1 - 0.9917^2}}$$

$$\cong 0.7023, \text{ or } 70.23\%$$

و

$$r_{YX_2 \cdot X_1} = \frac{r_{YX_2} - r_{YX_1} r_{X_1 X_2}}{\sqrt{1 - r^2_{X_1 X_2}}\sqrt{1 - r^2_{YX_1}}} = \frac{0.9917 - (0.9854)(0.9725)}{\sqrt{1 - 0.9725^2}\sqrt{1 - 0.9854^2}} \cong 0.8434, \text{ or } 84.34\%.$$

وعليه ، فإن X_2 أكثر أهمية من X_1 فى تفسير التغير فى Y .

مثال ٦ : يمكن تلخيص النتائج الكلية لمثال الحنطة – السماد – المبيد كالآتى :

$$\hat{Y} = 31.98 + 0.65X_1 + 1.10X_2$$

قيم — t : (2.70) (4.11)

$$R^2 = 0.992 \qquad \bar{R}^2 = 0.989 \qquad F_{2,7} = 413.17$$

$$r_{YX_1 \cdot X_2} = 0.70 \qquad r_{YX_2 \cdot X_1} = 0.84$$

وبالرغم من الحصول على النتائج عادة باستخدام الكمبيوتر ، إلا أنه من المهم القيام بحل المسألة « يدوياً » كما فعلنا لكى نفهم خطوات الحل بوضوح . وتعرض المسألة ٧ – ٢٢ عينة برنامج – كمبيوتر كامل يشرح بالكامل كيفية استخدام Statistical Package for the Social Sciences) SPSS وهو أكثر برامج الكمبيوتر شيوعاً فى الاستخدام) ، لانحدار متعدد فى ثلاث متغيرات .

مسائل محلولة

النموذج الخطى ذو المتغيرات الثلاثة :

٧ – ١ (أ) اكتب معادلة نموذج الانحدار الخطى المتعدد لحالة متغيرين مستقلين أو مفسرين وحالة k متغير مستقل أو مفسر
(ب) اذكر فروض النموذج الخطى للانحدار المتعدد .

(أ) فى حالة متغيرين مستقلين أو مفسرين ، المعادلة هى :

$$Y_i = b_0 + b_1X_{1i} + b_2X_{2i} + u_i \qquad (٧ - ١)$$

وفى حالة k متغير مستقل أو مفسر ، المعادلة هى :

$$Y_i = b_0 + b_1X_{1i} + b_2X_{2i} + \cdots + b_kX_{ki} + u_i$$

حيث تمثل X_{2i} ، على سبيل المثال . المشاهدة التى ترتيبها i للمتغير المستقل X_2 .

(ب) الفروض الخمسة الأول لنموذج الانحدار الخطى المتعدد هى نفس فروض نموذج الانحدار البسيط OLS (انظر المسألة ٦ – ٤) . أى أن الفروض الثلاثة الأول يمكن تلخيصها على النحو $u_i \sim N(0, \sigma$. الفرض الرابع هو $E(u_iu_j) = 0$ عند $i \neq j$ ؛ والفرض الخامس هو $E(X_iu_i) = 0$. الفرض الإضافى الوحيد المطلوب لنموذج الانحدار الخطى المتعدد OLS هو أنه لا توجد علاقة خطية تامة بين المتغيرات المستقلة $X_1, X_2, \ldots$ لأنه لو كان بين اثنين أو أكثر من المتغيرات المستقلة ارتباط خطى تام ، لاستحال حساب تقديرات معالم OLS لأن مجموعة المعادلات الطبيعية سوف تشتمل على معادلتين أو أكثر ليست مستقلة . أما إذا كان هناك ارتباط خطى كبير وليس تاماً بين اثنين أو أكثر من المتغيرات المفسرة ، فإنه يمكن تقدير معالم OLS ، ولكن لا يمكن عزل تأثير كل من المتغيرات المستقلة ذات الارتباط الخطى الكبير فيما بينها (أنظر قسم ٩ – ١)

٧ – ٢ باستخدام طريقة OLS فى حالة متغيرين مستقلين أو مفسرين ، اشتق (أ) المعادلة الطبيعية (٧ – ٢) ، (ب) المعادلة الطبيعية (٧ – ٣) ، (ج) المعادلة الطبيعية (٧ – ٤) . (القارئ غير الملم بالتفاضل يمكنه أن يتخطى هذه المسألة) .

(أ) تشتق المعادلة الطبيعية (٧ – ٢) بإيجاد النهاية الصغرى للمقدار $\sum e_i^2$ بالنسبة إلى $\hat{b}_0$:

$$\frac{\partial e_i^2}{\partial \hat{b}_0} = \frac{\partial \sum (Y_i - \hat{b}_0 - \hat{b}_1X_{1i} - \hat{b}_2X_{2i})^2}{\partial \hat{b}_0} = 0$$

$$-2\sum (Y_i - \hat{b}_0 - \hat{b}_1X_{1i} - \hat{b}_2X_{2i}) = 0 \qquad (٧ - ٢)$$

$$\sum Y_i = n\hat{b}_0 \pm \hat{b}_1 \sum X_{1i} \pm \hat{b}_2 \sum X_{2i}$$

(ب) وتشتق المعادلة الطبيعية (٧ – ٣) بإيجاد النهاية الصغرى للمقدار $\sum e_i^2$ بالنسبة إلى $\hat{b}_1$:

$$\frac{\partial \sum e_i^2}{\partial \hat{b}_1} = \frac{\partial \sum (Y_i - \hat{b}_0 - \hat{b}_1 X_{1i} - \hat{b}_2 X_{2i})^2}{\partial \hat{b}_1} = 0$$

$$-2X_{1i}\sum (Y_i - \hat{b}_0 - \hat{b}_1 X_{1i} - \hat{b}_2 X_{2i}) = 0 \qquad (٧ - ٣)$$

$$\sum X_{1i} Y_i = \hat{b}_0 \sum X_{1i} + \hat{b}_1 \sum X_{1i}^2 + \hat{b}_2 \sum X_{1i} X_{2i}$$

(ج) وتشتق المعادلة الطبيعية (٧ - ٤) بإيجاد النهاية الصغرى للمقدار $\sum e_i^2$ بالنسبة إلى $\hat{b}_2$:

$$\frac{\partial \sum e_i^2}{\partial \hat{b}_2} = \frac{\partial \sum (Y_i - \hat{b}_0 - \hat{b}_1 X_{1i} - \hat{b}_2 X_{2i})^2}{\partial \hat{b}_2} = 0$$

$$-2X_{2i}\sum (Y_i - \hat{b}_0 - \hat{b}_1 X_{1i} - \hat{b}_2 X_{2i}) = 0 \qquad (٧ - ٤)$$

$$\sum X_{2i} Y_i = \hat{b}_0 \sum X_{2i} + \hat{b}_1 \sum X_{1i} X_{2i} + \hat{b}_2 \sum X_{2i}^2$$

٧ - ٣ بالنسبة لنموذج الانحدار الخطى المتعدد ذى المتغيرين المستقلين ، (أ) اشتق المعادلات الطبيعية باستخدام الانحرافات (**ارشاد :** ابدأ باشتقاق تعبير $\hat{y}_i$ ؛ يمكن للقارىء غير الملم بالتفاضل أن يتخطى هذا الجزء من المسألة) . (ب) كيف يمكن اشتقاق المعادلات ٧ - ٥ ، ٧ - ٦ ، ٧ - ٧ ، لإيجاد $\hat{b}_1$ ، $\hat{b}_2$ ، $\hat{b}_0$ ؟

(أ)

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_{1i} + \hat{b}_2 X_{2i}$$

$$\bar{Y} = \hat{b}_0 + \hat{b}_1 \bar{X}_1 + \hat{b}_2 \bar{X}_2$$

بالطرح ، نحصل على

$$\hat{y}_i = \hat{Y}_i - \bar{Y} = \hat{b}_1 x_{1i} + \hat{b}_2 x_{2i}$$

وعليه

$$e_i = y_i - \hat{y}_i = y_i - \hat{b}_1 x_{1i} - \hat{b}_2 x_{2i}$$

$$\sum e_i^2 = \sum (y_i - \hat{y}_i)^2 = \sum (y_i - \hat{b}_1 x_{1i} - \hat{b}_2 x_{2i})^2$$

$$\frac{\partial \sum e_i^2}{\partial \hat{b}_1} = \frac{\partial \sum (y_i - \hat{b}_1 x_{1i} - \hat{b}_2 x_{2i})^2}{\partial \hat{b}_1} = 0$$

$$-2x_{1i}\sum (y_i - \hat{b}_1 x_{1i} - \hat{b}_2 x_{2i}) = 0$$

$$\sum x_{1i} y_i = \hat{b}_1 \sum x_{1i}^2 + \hat{b}_2 \sum x_{1i} x_{2i} \qquad (٧ - ١٦)$$

$$\frac{\partial \sum e_i^2}{\partial \hat{b}_2} = \frac{\partial \sum (y_i - \hat{b}_1 x_{1i} - \hat{b}_2 x_{2i})^2}{\partial \hat{b}_2} = 0$$

$$-2x_{2i}\sum (y_i - \hat{b}_1 x_{1i} - \hat{b}_2 x_{2i}) = 0$$

$$\sum x_{2i} y_i = \hat{b}_1 \sum x_{1i} x_{2i} + \hat{b}_2 \sum x_{2i}^2 \qquad (٧ - ١٧)$$

(ب) المعادلات (٧ - ٥) ، (٧ - ٦) لحساب $\hat{b}_1$ و $\hat{b}_2$ على الترتيب ، يتم الحصول عليها بحل معادلات (٧ - ١٦) ، (٧ - ١٧) آنياً . ويمكن دائماً حساب $\hat{b}_1$ و $\hat{b}_2$ ، إلا إذا كانت هناك علاقة خطية تامة بين X_1 و X_2 أو كان عدد المشاهدات عن كل متغير فى النموذج 3 أو أقل . ويمكن حساب المعلمة $\hat{b}_0$ بالتعويض فى معادلة (٧ - ٧) بقيم $\hat{b}_1$ و $\hat{b}_2$ (المحسوبة باستخدام معادلات (٧ - ٥) ، (٧ - ٦)) رقيم $\bar{Y}$ و $\bar{X}_1$ و $\bar{X}_2$ (المحسوبة من معطيات المسألة) .

٧ — ٤ بالنسبة لتحليل الانحدار المتعدد ذى المتغيرين المستقلين بين معنى (أ) b_0 (ب) b_1 ، (جـ) b_2 (د) هل $\hat{b}_0$ و $\hat{b}_1$ و $\hat{b}_2$ BLUE؟

(أ) المعلمة b_0 هى الحد الثابت أو مقطع الانحدار وتعطى قيمة المتغير Y_i ، عندما $X_{1i} = X_{2i} = 0$.

(ب) تقيس المعلمة b_1 التغير فى Y لكل وحدة تغير فى X_1 مع إبقاء X_2 ثابتة . ومعلمة الميل b_1 هى معامل انحدار جزئى لأنها تناظر المشتقة الجزئية للمتغير Y بالنسبة إلى X_1 ، أى $\partial Y/\partial X_1$.

(جـ) تقيس المعلمة b_2 التغير فى Y لكل وحدة تغير فى X_2 مع إبقاء X_1 ثابتة . ومعلمة الميل b_2 هى المعامل الجزئى الثانى للانحدار لأنها تناظر المشتقة الجزئية للمتغير Y بالنسبة إلى X_2 ، أى $\partial Y/\partial X_2$.

(د) حيث أنه يتم الحصول على $\hat{b}_0$ و$\hat{b}_1$ وb_2 بطريقة OLS ، فإنها أيضاً أفضل مقدرات خطية غير متحيزة (BLUE ، انظر قسم ٦ — ٥) . أى أن $E(\hat{b}_0)=b_0$ ، $E(\hat{b}_1) = b_1$ ، $E(\hat{b}_2)=b_2$ و $S^2_{\hat{b}0}$ ، $S^2_{\hat{b}1}$ ، $S^2_{\hat{b}2}$ أصغر منها لأى مقدرات خطية غير متحيزة أخرى ولما كان إثبات هذه الخصائص يمثل عبئاً ثقيلا بدون استخدام جبر المصفوفات لذا لا نتناولها هنا .

٧ — ٥ جدول ٧ — ٣ هو امتداد لجدول ٦ — ١١ ويعطى دخل الفرد الحقيقى بآلاف الدولارات Y ، مع نسبة القوة العاملة فى الزراعة ، X_1 ومتوسط سنوات التعليم للسكان فوق سن ٢٥ سنة ، X_2 لعدد 15 دولة متقدمة فى ١٩٨١ . (أ) أوجد معادلة انحدار المربعات الصغرى للمتغير Y على X_1 وX_2 (ب) فسر النتائج فى (أ) وقارنها بنتائج المسألة ٦ — ٣٠ .

جدول ٧ — ٣ دخل الفرد ، القوة العاملة فى الزراعة ، وسنوات التعليم

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Y	6	8	8	7	7	12	9	8	9	10	10	11	9	10	11
X_1	9	10	8	7	10	4	5	5	6	8	7	4	9	5	8
X_2	8	13	11	10	12	16	10	10	12	14	12	16	14	10	12

جدول ٧ — ٤ مسودة لتقدير المعالم لبيانات جدول ٧ — ٣

n	Y	X_1	X_2	y	x_1	x_2	x_1y	x_2y	x_1x_2	x_1^2	x_2^2
1	6	9	8	−3	2	−4	−6	12	−8	4	16
2	8	10	13	−1	3	1	−3	−1	3	9	1
3	8	8	11	−1	1	−1	−1	1	−1	1	1
4	7	٧	10	−2	0	−2	0	4	0	0	4
5	7	10	12	−2	3	0	−6	0	0	9	0
6	12	4	16	3	−3	4	−9	12	−12	9	16
7	9	5	10	0	−2	−2	0	0	4	4	4
8	8	5	10	−1	−2	−2	2	2	4	4	4
9	9	6	12	0	−1	0	0	0	0	1	0
10	10	8	14	1	1	2	1	2	2	1	4
11	10	7	12	1	0	0	0	0	0	0	0
12	11	4	16	2	−3	4	−6	8	−12	9	16
13	9	9	14	0	2	2	0	0	4	4	4
14	10	5	10	1	−2	−2	−2	−2	4	4	4
15	11	8	12	2	1	0	2	0	0	1	0
$n = 15$	$\sum Y = 135$ $\bar{Y} = 9$	$\sum X_1 = 105$ $\bar{X}_1 = 7$	$\sum X_2 = 180$ $\bar{X}_2 = 12$	$\sum y = 0$	$\sum x_1 = 0$	$\sum x_2 = 0$	$\sum x_1 y = -28$	$\sum x_2 y = 38$	$\sum x_1 x_2 = -12$	$\sum x_1^2 = 60$	$\sum x_2^2 = 74$

(أ) يبين جدول ٧ – ٤ الحسابات اللازمة لتقدير معالم معادلة انحدار OLS للمتغير Y على المتغيرين X_1 و X_2 .

$$\hat{b}_1 = \frac{\left(\sum x_1 y\right)\left(\sum x_2^2\right) - \left(\sum x_2 y\right)\left(\sum x_1 x_2\right)}{\left(\sum x_1^2\right)\left(\sum x_2^2\right) - \left(\sum x_1 x_2\right)^2} = \frac{(-28)(74) - (38)(-12)}{(60)(74) - (-12)^2} = \frac{-2{,}072 + 456}{4{,}440 - 144} \cong -0.38$$

$$\hat{b}_2 = \frac{\left(\sum x_2 y\right)\left(\sum x_1^2\right) - \left(\sum x_1 y\right)\left(\sum x_1 x_2\right)}{\left(\sum x_1^2\right)\left(\sum x_2^2\right) - \left(\sum x_1 x_2\right)^2} = \frac{(38)(60) - (-28)(-12)}{(60)(74) - (-12)^2} = \frac{2{,}280 - 336}{4{,}440 - 144} \cong 0.45$$

$$\hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X}_1 - \hat{b}_2 \bar{X}_2 \cong 9 - (-0.38)(7) - (0.45)(12) = 9 + 2.66 - 5.40 \cong 6.26$$

وعليه فمعادلة OLS لانحدار Y على X_1 و X_2 هي :

$$\hat{Y}_i = 6.26 - 0.38\, X_{1i} + 0.45\, X_{2i}$$

(ب) تشير معادلة انحدار OLS المقدرة على أن مستوى دخل الفرد الحقيقي Y ، يرتبط عكسياً مع نسبة القوة العاملة فى الزراعة X_1 ، وطردياً مع عدد سنوات التعليم للسكان فوق سن 25 (كما قد يكون متوقعاً). بالتحديد تشير $\hat{b}_1$، إلى أن نقص نسبة القوة العاملة فى الزراعة بمقدار 1% من إجمالى القوة العاملة سوف يصاحبه زيادة قدرها 380 دولارا أمريكياً فى دخل الفرد مع تثبيت X_2. ولكن ، زيادة سنة واحدة فى سنوات التعليم للسكان فوق سن 25 سنة يصاحبها زيادة فى دخل الفرد قدرها 450 دولاراً أمريكياً ، مع تثبيت X_1 . وعند $X_{1i} = X_{2i} = 0,\ \hat{Y}_i = \hat{b}_0 = 6.26$. وبقدر ما اتضح أن X_2 معنوية إحصائياً (انظر المسألة ٧ – ٨ (ب)) ، وبالتالى يجب أن تدخل فى علاقة الانحدار ، فقد اتضح أيضاً أن $\hat{b}_1 \cong -0.47$ السابق إيجادها فى تمرين ٦ – ٣٠ لا تكون تقديراً موثوقاً للمعلمة b_1 .

اختبارات معنوية تقديرات المعالم :

٧ – ٦ عرف (أ) σ_u^2 و s^2 ، (ب) تباين $\hat{b}_1$ وتباين $\hat{b}_2$ ، (جـ) $s_{\hat{b}_1}^2$ و $s_{\hat{b}_2}^2$ ، (د) $s_{\hat{b}_1}$ و $s_{\hat{b}_2}$. (هـ) لماذا لا تكون b_0 عادة موضع اهتمام أساسى ؟

(أ) σ_u^2 هو تباين حدا الخطأ فى العلاقة الحقيقية بين X_{1i} و X_{2i} و Y_i . ولكن $s^2 = \hat{\sigma}_u^2 = \sum e_i^2/(n-k)$ هى تباين البواقى وهى تقدير غير متحيز للتباين غير المعلوم σ_u^2 . k هى عدد المعالم المقدرة . فى حالة الانحدار المتعدد ذى المتغيرين ، $k = 3$. وعليه $n - k = n - 3 = \text{df}$.

(ب)

$$\operatorname{Var}\hat{b}_1 = \sigma_u^2 \frac{\sum x_2^2}{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2}$$

بينما

$$\operatorname{Var}\hat{b}_2 = \sigma_u^2 \frac{\sum x_1^2}{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2}$$

إن تباينى $\hat{b}_1$ و $\hat{b}_2$ (أو تقديراتها) مطلوبة لاختبار الفروض وتكوين فترات الثقة لكل من b_1 و b_2 .

(جـ)

$$s_{\hat{b}_1}^2 = s^2 \frac{\sum x_2^2}{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2} = \frac{\sum e_i^2}{n-k} \cdot \frac{\sum x_2^2}{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2}$$

$$s_{\hat{b}_2}^2 = s^2 \frac{\sum x_1^2}{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2} = \frac{\sum e_i^2}{n-k} \cdot \frac{\sum x_1^2}{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2}$$

$s_{\hat{b}_1}^2$ و $s_{\hat{b}_2}^2$ هما على الترتيب ، تقديران غير متحيزين لتباين b_1 وتباين b_2 غير المعلومين حيث أن σ_u^2 غير معلومة .

(د) $s_{\hat{b}_1} = \sqrt{s_{\hat{b}_1}^2}$ و $s_{\hat{b}_2} = \sqrt{s_{\hat{b}_2}^2}$. $s_{\hat{b}_1}$ و $s_{\hat{b}_2}$ هما ، على الترتيب ، الانحراف المعيارى لكل من $\hat{b}_1$ و $\hat{b}_2$ ويسميان بالأخطاء المعيارية .

(ه) ما لم تتوفر مشاهدات كافية بالقرب من $X_{1i} = X_{2i} = 0$ فإن معلمة المقطع b_0 لا تكون عادة ذات أهمية أساسية ويمكن حذف اختبار معنوية الإحصائية الخاص بها ومعادلة (٧ – ١٨) لتباين $\hat{b}_0$ معادلة معقدة فى الحساب ولهذا السبب أيضاً فن النادر أن تذكر أو تستخدم :

$$\operatorname{Var} \hat{b}_0 = \sigma_u^2 \cdot \frac{\sum X_1^2 \sum X_2^2 - \left(\sum X_1 X_2\right)^2}{n\left[\sum X_1^2 X_2^2 - \left(\sum X_1 X_2\right)^2\right] - \sum X_1\left(\sum X_1 \sum X_2^2 - \sum X_2 \sum X_1 X_2\right) + \sum X_2\left(\sum X_1 \sum X_1 X_2 - \sum X_2 \sum X_1^2\right)} \quad (٧ - ١٨)$$

ومع ذلك ، ترد $S_{\hat{b}_0}^2$ أحياناً فى نتائج الكمبيوتر ، ويمكن إجراء الاختبارات الإحصائية لمعنوية b_0 بسهولة .

٧ – ٧ من بيانات جدول ٧ – ٣ ، أوجد (أ) S^2 ، (ب) $s_{\hat{b}_1}^2$ و $s_{\hat{b}_1}$ ، (ج) $s_{\hat{b}_2}^2$ و $s_{\hat{b}_2}$.

(أ) الحسابات اللازمة لإيجاد S^2 موضحة فى جدول ٧ – ٥ ، وهو امتداد لجدول ٧ – ٤ . وقد تم الحصول على قيم $\hat{Y}_i$ بالتعويض بقيم X_{1i} و X_{2i} فى معادلة انحدار OLS المقدرة السابق إيجادها فى المسألة ٧ – ٥ (أ) :

$$s^2 = \hat{\sigma}_u^2 = \frac{\sum e_i^2}{n-k} = \frac{12.2730}{15-3} \cong 1.02$$

جدول ٧ – ٥ انحدار دخل الفرد : حسابات اختبار معنوية العالم

الدولة	Y	X_1	X_2	$\hat{Y}$	e	e^2
1	6	9	8	6.44	− 0.44	0.1936
2	8	10	13	8.31	− 0.31	0.0961
3	8	8	11	8.17	− 0.17	0.0289
4	7	7	10	8.10	− 1.10	1.2100
5	7	10	12	7.86	− 0.86	0.7396
6	12	4	16	11.94	0.06	0.0036
7	9	5	10	8.86	0.14	0.0196
8	8	5	10	8.86	− 0.86	0.7396
9	9	6	12	9.38	− 0.38	0.1444
10	10	8	14	9.52	0.48	0.2304
11	10	7	12	9.00	1.00	1.0000
12	11	4	16	11.94	− 0.94	0.8836
13	9	9	14	9.14	− 0.14	0.0196
14	10	5	10	8.86	1.14	1.2996
15	11	8	12	8.62	2.38	5.6644
$n = 15$					$\sum e = 0$	$\sum e^2 = 12.2730$

(ب) باستخدام قيمة S^2 السابق إيجادها فى (أ) وقيم جدول ٧ – ٤ ، نحصل على :

$$s^2_{\hat{b}_1} = s^2 \frac{\sum x_2^2}{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2} \cong 1.02 \frac{74}{(60)(74) - (-12)^2} \cong 0.02$$

$$s_{\hat{b}_1} \cong \sqrt{0.02} \cong 0.14$$

(جـ)

$$s^2_{\hat{b}_2} = s^2 \frac{\sum x_1^2}{\sum x_1^2 \sum x_2^2 - \left(\sum x_1 x_2\right)^2} \cong 1.02 \frac{60}{(60)(74) - (-12)^2} \cong 0.01$$

$$s_{\hat{b}_2} \cong \sqrt{0.01} \cong 0.10$$

٧ – ٨ اختبر عند مستوى معنوية 5% كل من (أ) b_1 (ب) b_2 فى المسألة ٧ – ٥ (أ) .

(أ)

$$t_1 = \frac{\hat{b}_1 - b_1}{s_{\hat{b}_1}} = \frac{-0.38 - 0}{0.14} \cong 2.71$$

وحيث أن قيمة t_1 المطلقة تتجاوز القيمة الجدولية $t = 2.179$ (من ملحق ٥) بمستوى معنوية 5% (اختبار ذو ذيلين) ، $n - k = 15 - 3 = 12\text{df}$ ، فإننا نستنتج أن b_1 معنوية إحصائياً عند مستوى معنوية 5% (أى أننالا نستطيع أن نرفض H_1 ، بأن $b_1 \neq 0$) .

(ب)

$$t_2 = \frac{\hat{b}_2 - b_2}{s_{\hat{b}_2}} \cong \frac{0.45 - 0}{0.10} = 4.50$$

أى أن b_2 معنوية إحصائياً عند مستوى معنوية 5% (وأيضاً 1%) (أى أنه لا يمكن رفض H_1 بأن $b_2 \neq 0$) .

٧ – ٩ كون فترة الثقة 95% لكل من (أ) b_1 (ب) b_2 فى المسألة ٧ – ٥ (أ) .

(أ) فترة الثقة 95% للمعلمة b_1 :

$$b_1 = \hat{b}_1 \pm 2.179\, s_{\hat{b}_1} = -0.38 \pm 2.179(0.14) = -0.38 \pm 0.31$$

أى أن b_1 بين -0.69 ، -0.07 (أى $-0.69 \leq b_1 \leq -0.07$) (بدرجة ثقة 95% .

(ب) فترة الثقة 95% للمعلمة b_2 :

$$b_2 = \hat{b}_2 \pm 2.179\, s_{\hat{b}_2} = 0.45 \pm 2.179(0.10) = 0.45 \pm 0.22$$

أى أن b_2 بين 0.23 ، 0.67 (أى $0.23 \leq b_2 \leq 0.67$) بدرجة ثقة 95%

معامل التحديد المتعدد :

٧ – ١٠ بدءاً باستخدام $R^2 = 1 - \sum e_i^2 / \sum y_i^2$ | اشتق | $R^2 = (\hat{b}_1 \sum y x_1 + \hat{b}_2 \sum y x_2)/\sum y_i^2$

(إرشاد : ابدأ بتبيان أن $\sum e_i^2 = \sum y_i^2 - \hat{b}_1 \sum y x_1 - \hat{b}_2 \sum y x_2$ يمكن القارئ غير الملم بالتفاضل أن يتخطى هذه المسألة) .

$$\sum e_i^2 = \sum e_i (y_i - \hat{y}_i) = \sum e_i (y_i - \hat{b}_1 x_{1i} - \hat{b}_2 x_{2i}) = \sum e_i y_i - \hat{b}_1 \sum e_i x_{1i} - \hat{b}_2 \sum e_i x_{2i}$$

ولكن من طريقة OLS وجدنا

$$\frac{\partial \sum e_i^2}{\partial \sum \hat{b}_1} = -\sum e_i x_{1i} = 0 \quad \text{and} \quad \sum e_i x_{1i} = 0$$

$$\frac{\partial \sum e_i^2}{\partial \sum \hat{b}_2} = -\sum e_i x_{2i} = 0 \quad \text{و} \quad \sum e_i x_{2i} = 0$$

وبالتالى

$$\sum e_i^2 = \sum e_i y_i = \sum (y_i - \hat{y}_i) y_i = \sum y_i (y_i - \hat{b}_1 \sum y_i x_{1i} - \hat{b}_2 \sum y_i x_{2i})$$
$$= \sum y_i^2 - \hat{b}_1 \sum y_i x_{1i} - \hat{b}_2 \sum y_i x_{2i}$$

بالتعويض فى معادلة R^2

$$R^2 = 1 - \frac{\sum e_i^2}{\sum y_i^2} = 1 - \frac{\sum y_i^2 - \hat{b}_1 \sum y_i x_{1i} - \hat{b}_2 \sum y_i x_{2i}}{\sum y_i^2} = \frac{\hat{b}_1 \sum y_i x_{1i} + \hat{b}_2 \sum y_i x_{2i}}{\sum y_i^2}$$

أو بحذف i للتبسيط ، نحصل على (كما فى قسم ٧ – ٣)

$$R^2 = \frac{\hat{b}_1 \sum y x_1 + \hat{b}_2 \sum y x_2}{\sum y^2}$$

٧ – ١١ أوجد R^2 من معادلة انحدار OLS المقدرة فى المسألة ٧ – ٥ (أ) ، باستخدام (أ) $R^2 = \sum \hat{y}_i^2 / \sum y_i^2$ (ب) $R^2 = 1 - \sum e_i^2 / \sum y_i^2$ (جـ) $R^2 = (\hat{b}_1 \sum y x_1 + \hat{b}_2 \sum y x_2) / \sum y_i^2$

(أ) نعرف من المسألة ٦ – ٢٠ أن

$$\sum y_i^2 = \sum \hat{y}_i^2 + \sum e_i^2 \qquad \text{so that} \qquad \sum \hat{y}_i^2 = \sum y_i^2 - \sum e_i^2$$

حيث $\sum y_i^2 = 40$ (بتربيع وجمع قيم y_i من جدول ٧ – ٤) $\sum e_i^2 = 12.2730$ (من جدول ٧ – ٥) ، وبالتالى $\sum \hat{y}_i^2 = 40 - 12.2730 = 27.7270$. $R^2 = \sum \hat{y}_i^2 / \sum y_i^2 = 27.7270/40 \cong 0.6932$, or 69.32% .

(ب) باستخدام $\sum e_i^2 = 12.2730$ و $\sum y_i^2 = 40$ نحصل على $R^2 = 1 - \sum e_i^2/\sum y_i^2 = 1 - 12.2730/40 \cong 0.6932$ أو 69.32% 69.32% كما فى (أ)

(جـ) باستخدام $b_1 = 0.38$ و $b_2 = 0.45$ (السابق إيجادها فى مسالة ٧ – ٥ (أ) ، $\sum y x_1 = -28$ و $\sum y x_2 = 38$ (من جدول ٧ – ٤) ، $\sum y_i^2 = 40$ نحصل على

$$R^2 = \frac{\hat{b}_1 \sum y x_1 + \hat{b}_2 \sum y x_2}{\sum y^2} = \frac{(-0.38)(-28) + (0.45)(38)}{40} \cong \frac{27.74}{40} = 0.6935, \text{ or } 69.35\%$$

وتختلف قيمة R^2 هذه قليلا عن تلك السابق إيجادها فى (أ) ، (ب) كنتيجة لأخطاء التقريب .

٧ – ١٢ من $R^2 = 1 - (\sum e_i^2 / \sum y_i^2)$ اشتق $\bar{R}^2$ (ب) ما هو المدى لقيم $\bar{R}^2$ (ارشاد بالنسبة لجزء (أ) : ابدأ بالتشابه بين $\sum e_i^2$ وتباين $\sum y_i^2$ وتباين Y) .

(أ) صعوبة R^2 (غير المعدلة) أنها لا تأخذ فى الاعتبار درجات الحرية ولكن $e = s^2 = \sum e_i^2/(n-k)$ حيث $n - k = \text{df}$ درجات الحرية و $\text{var}\, Y = \sum (Y_i - \bar{Y})^2/(n-1)$ حيث $n - 1 = \text{df}$. وبالتالى $\sum e_i^2 = s^2(n-k)$ و $\sum (Y_i - \bar{Y})^2 = \sum y_i^2 = \text{var}\, Y(n-1)$ فتكون

$$R^2 = 1 - \frac{\sum e_i^2}{\sum y_i^2} = 1 - \frac{s^2(n-k)}{\text{var}\, Y(n-1)}$$

وعليه $1 - R^2 = (s^2/\text{var}\, Y)(n-k)/(n-1)$ ولكن $1 - \bar{R}^2 = s^2/\text{var}\, Y$ فتكون

$$1 - R^2 = (1 - \bar{R}^2)\frac{(n-k)}{(n-1)}$$

وبالحل لإيجاد R^2 ، نحصل على

$$\bar{R}^2 = 1 - (1 - R^2)\frac{(n-1)}{(n-k)} \qquad (٧ - ١٢)$$

(ب) عندما $k = 1$ ، $(n-1)/(n-k) = 1$ تكون $R^2 = \bar{R}^2$

عندما $k > 1$ ، $(n-1)/(n-k) > 1$ تكون $R^2 > \bar{R}^2$

عندما تكون n كبيرة ، لقيمة معينة k ، تكون $(n-k)/(n-1)$ قريبة من الوحدة ، ولن تختلف $\bar{R}^2$ عن R^2 كثيراً . عندما تكون n صغيرة وتكون k كبيرة بالنسبة إلى n ، فإن $\bar{R}^2$ سوف تكون أصغر كثيرا من R^2 وقد تكون $\bar{R}^2$ (سالبة بالرغم من أن $0 \leq R^2 \leq 1$) . أنظر المسائل من ٧ – ٢٩ إلى ٧ – ٣٢ .

٧ – ١٣ (أ) أوجد $\bar{R}^2$ بالنسبة لمعادلة انحدار OLS المقدرة فى مسألة ٧ – ٥ (أ) .

(ب) كيف تقارن $\bar{R}^2$ المحسوبة فى (أ) مع R^2 فى مسألة ٧ – ١١ (أ) ، فى مسألة ٦ – ٣١ (جـ) ؟

(أ) باستخدام $R^2 = 0.6932$ السابق إيجادها فى المسألة ٧ – ١١ (ب) ، نحصل على

$$\bar{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-k} = 1 - (1 - 0.6932)\frac{15-1}{15-3} \cong 0.6410$$

(ب) $R^2 = 0.33$ فى حالة الانحدار البسيط ، باستخدام نسبة القوة العاملة فى الزراعة X_1 ، كتغير مستقل وحيد (انظر المسألة ٦ – ٣١ (جـ)) . $R^2 = 0.69$ بعد إضافة سنوات التعليم للسكان فوق سن 25 سنة ، X_2 كتغير مستقل ثان . ولكن ، عندما نأخذ فى الاعتبار حقيقة أن إضافة X_2 يقلل درجات الحرية بمقدار 1 $n - k = 15 - 2 = 13$ فى حالة الانحدار البسيط للمتغير Y على X_1 ، إلى $n - k = 15 - 3 = 12$ فى الانحدار المتعدد للمتغير Y على X_1 و X_2) ، فإن $\bar{R}^2$ تنخفض إلى 0.64

وحقيقة أن b_2 وجدت معنوية إحصائياً (فى المسألة ٧ – ٨ (ب)) $R^2 = \bar{R}^2 = 0.33$ فى حالة الانحدار البسيط المتغير Y على X_1 ترتفع إلى $\bar{R}^2 = 0.64$ فى حالة الانحدار المتعدد للمتغير Y على X_1 و X_2 يبرر الإبقاء على X_2 كتغير مستقل إضافى فى معادلة الانحدار .

٧ – ١٤ (أ) كيف يمكن إيجاد $\sum e_i^2$ (المطلوبة لإجراء اختبارات المعنوية) بدون إيجاد $\hat{Y}_i$ أولا ؟ (ب) أوجد $\sum e_i^2$ لبيانات جدول ٧ – ٣ بدون إيجاد $\hat{Y}_i$ (جدول ٧ – ٥) .

(أ) باستخدام القيم المقدرة لكل من b_1 و b_2 وكذلك Σyx_1 ، Σyx_2 و Σy^2 نحصل أولا على :

$$R^2 = \frac{\hat{b}_1\sum yx_1 + \hat{b}_2\sum yx_2}{\sum y^2}$$

وبالتالى $R^2 = 1 - (\sum e_i^2/\sum y_i^2)$, so فتكون $\sum e_i^2 = (1 - R^2)\sum y_i^2$. وهذه الطريقة لإيجاد $\sum e_i^2$ تتضمن حسابات أقل عن استخدام Y_i (فالحسابات الوحيدة الإضافية بجانب تلك المطلوبة لتقدير b_1 و b_2 هى $\sum y_i^2$) .

(ب) من قيمة $R^2 = 0.6935$ السابق إيجادها فى المسألة ٧ – ١١ (جـ) (التى تستخدم فقط القيم المقدرة لكل من $\hat{b}_1$ و $\hat{b}_2$ السبق إيجادها فى مسألة ٧ – ٥ (أ) والقيم المحسوبة فى جدول ٧ – ٤) . ومن قيمة $\sum y_i^2 = 40$ من المسألة ٧ – ١١ (أ) ، نحصل على

$$\sum e_i^2 = (1 - R^2)\sum y_i^2 = (1 - 0.6935)(40) = 12.26$$

قارن هذه بقيمة $\sum e_i^2 = 12.2730$ السابق إيجادها فى جدول ٧ – ٥ . (الفرق الصغير فى قيمتى $\sum e_i^2$ اللتين حصلنا عليهما باستخدام الطريقتين راجع إلى أخطاء التقريب) . لاحظ ، أن إيجاد $\sum e_i^2$ بالطريقة السابقة يلغى تماماً الحاجة إلى جدول ٧ – ٥ .

اختبار المعنوية الإجمالية للانحدار :

٧ – ١٥ أذكر الفرض العدمى والفرض البديل لاختبار معنوية الانحدار ككل . (ب) كيف تختبر المعنوية الكلية للانحدار ؟ ما هو منطق هذا الاختبار ؟ (جـ) أعط صيغة التباين المفسر ، التباين غير المفسر أو تباين البواقى .

(أ) يشير اختبار المعنوية الإجمالية للانحدار إلى اختبار الفرض أ أكل المتغيرات المستقلة لا تساعد على تفسير التغير فى المتغير التابع حول وسطه . وبشكل محدد ، الفرض العدمى هو :

$$H_0: \quad b_1 = b_2 = \ldots = b_k = 0$$

مقابل الفرض البديل

ليست كل قيم b_i تساوى الصفر $H_1:$

(ب) تختبر المعنوية الكلية للانحدار بحساب النسبة F بين التباين المفسر والتباين غير المفسر أو تباين البواقى . وتوحى القيمة « المرتفعة » الإحصائية F بعلاقة معنوية بين المتغير التابع والمتغيرات المستقلة ، مؤدية إلى رفض الفرض العدمى بأن معاملات كل المتغيرات المفسرة كلها أصفار .

(جـ) التباين المفسر = $\sum(\hat{Y}_i - \bar{Y})^2/(k-1) = \text{RSS}/(k-1) = \sum \hat{y}_i^2/(k-1)$ حيث k عدد المعالم المقدرة (انظر قسم ٦ – ٤) . والتباين غير المفسر = $\sum e_i^2/(n-k) = \text{ESS}/(n-k) = \sum(Y_i - \hat{Y}_i)^2/(n-k)$.

٧ – ١٦ (أ) اعط صيغة إحصائية أو نسبة F المحسوبة لحالة الانحدار البسيط وللانحدار عند $n = 15$ ، $k = 3$ (ب) هل يمكن أن تكون F المحسوبة « كبيرة » ومع ذلك فكل المعالم المقدرة ليست معنوية إحصائياً ؟

(أ)
$$F_{1,n-2} = \frac{\sum \hat{y}_i^2/1}{\sum e_i^2/(n-2)}$$

حيث تشير رموز دليل F إلى عدد درجات الحرية فى البسط والمقام على الترتيب . فى حالة الانحدار البسيط ، $F_{1,n-2} = t_{n-2}^2$ لنفس مستوى الثقة . بالنسبة للانحدار المتعدد عند $n = 15$ ، $k = 3$ و $F_{2,12} = (\sum \hat{y}_i^2/2)/(\sum e_i^2/12)$.

(ب) من الممكن أن تكون F المحسوبة « كبيرة » وليس بين المعالم المحسوبة ما هو معنوى إحصائياً . وقد يحدث هذا عندما يكون هناك ارتباط مرتفع بين المتغيرات المستقلة بعضها البعض (انظر قسم ٩ – ٢) . وغالباً ما يكون اختبار F ذا فائدة محدودة لأنه من الممكن أن يرفض الفرض العدمى ، بصرف النظر عما إذا كان النموذج يشرح « جزءاً كبيراً » من التغير فى Y .

٧ – ١٧ (أ) أثبت أن $[\sum \hat{y}_i^2/(k-1)]/[\sum e_i^2/(n-k)] = [R^2/(k-1)]/[(1-R^2)/(n-k)]$.

(ب) على ضوء نتائج (أ) ، ما هى الطريقة البديلة للتعبير عن الفرض لاختبار المعنوية الكلية للانحدار ؟

(أ)
$$\frac{\sum \hat{y}_i^2/(k-1)}{\sum e_i^2/(n-k)} = \frac{\sum \hat{y}_i^2}{\sum e_i^2}\,\frac{n-k}{k-1} = \frac{\sum \hat{y}_i^2/\sum y_i^2}{\sum e_i^2/\sum y_i^2}\,\frac{n-k}{k-1} = \frac{R^2}{(1-R^2)}\,\frac{n-k}{k-1} = \frac{R^2/(k-1)}{(1-R^2)/(n-k)}$$

(ب) نسبة F ، كاختبار لمعنوية القدرة التفسيرية لكل المتغيرات المستقلة معاً ، تعادل تقريباً اختبار معنوية الإحصائية R^2 فإذا قبل الفرض البديل فإننا نتوقع أن تكون R^2 ، وبالتالى F ، « عالية » .

٧ – ١٨ اختبر عند مستوى معنوية 5% المعنوية الإجمالية لانحدار OLS المقدر فى المسألة ٧ – ٥ (أ) باستخدام (أ)

$$[\sum \hat{y}_i^2/(k-1)]/[\sum e_i^2/(n-k)]$$

(ب) $[R^2/(k-1)]/[(1-R^2)/(n-k)]$

(أ) باستخدام $\sum \hat{y}_i^2 = 27.727$ من المسألة ٧ – ١١ (أ) ، $\sum e_i^2 = 12.2730$ من جدول ٧ – ٥ ، نحصل على

$$F_{2,12} = \frac{27.727/2}{12.273/12} \cong 13.59$$

وحيث أن القيمة المحسوبة للنسبة F تفوق القيمة الجدولية $F=3.88$ عند مستوى معنوية 5% ودرجات حرية 2 ، 12 (انظر ملحق ٧) ، فإننا نقبل الفرض البديل بأنه ليست كل قيم b_i تساوى الصفر عند مستوى معنوية 5% ..

(ب) باستخدام $R^2 = 0.6932$ من تمرين ٧ – ١١ (ب) ، نحصل على

$$F_{2,12} = \frac{R^2/(k-1)}{(1-R^2)/(n-k)} = \frac{0.6932/2}{(1-0.6932)/12} \cong 13.54$$

ونقبل الفرض أن R^2 تختلف معنوياً عن الصفر عند مستوى معنوية 5% .

معاملات الارتباط الجزئى :

٧ – ١٩ (أ) كيف يمكن إبعاد تأثير X_2 عن كل من Y و X_1 عند إيجاد $r_{YX_1 \cdot X_2}$ ؟ (ب) ما هو المدى لقيم معاملات الارتباط الجزئى ؟ (جـ) ما هى إشارة معاملات الارتباط الجزئى ؟ (د) ما فائدة معاملات الارتباط الجزئى ؟

(أ) لإبعاد تأثير X_2 على Y ، فإننا نوجد انحدار Y على X_2 ، ونوجد البواقى $e_1 = Y^*$. ولإبعاد تأثير X_2 على X_1 فإننا نوجد انحدار X_1 و X_2 ونوجد البواقى $e_2 = X_1^*$. وفى هذه الحالة فان Y^* و X_1^* تمثلان التغير فى Y و X_1 على الترتيب ، الباقى بدون تفسير بعد إزاحة تفسير X_2 على كل من Y و X_1 . وبالتالى ، فمعامل الارتباط الجزئى ليس إلى معامل الارتباط البسيط بين البواقى Y^* و X_1^* (أى أن ، $r_{YX_1 \cdot X_2} = r_{Y^*X_1^*}$)

(ب) مدى معاملات الارتباط الجزئى هو من -1 إلى $+1$ (تماماً كما فى حالة معاملات الارتباط البسيط) . على سبيل المثال ، تشير $r_{YX_1 \cdot X_2} = -1$ إلى الحالة عندما توجد علاقة خطية تامة عكسية بين Y و X_1 بعد إزاحة التأثير المشترك للمتغير X_2 على كل من Y و X_1 . ولكن ، تشير $r_{YX_1 \cdot X_2} = 1$ إلى علاقة خطية تامة طردية صافية بين Y و X_1 . تشير $r_{YX_1 \cdot X_2} = 0$ إلى عدم وجود علاقة بين Y و X_1 بعد إزاحة تأثير X_2 على كل من Y و X_1 . وكنتيجة ، فإنه يمكن حذف X_1 من الانحدار .

(جـ) إشارة معامل الانحدار الجزئى هى نفس إشارة المعلمة المقدرة المناظرة . فمثلا ، بالنسبة لمعادلة الانحدار المقدرة $\hat{Y} = \hat{b}_0 + \hat{b}_1 X_1 + \hat{b}_2 X_2$، $r_{YX_1 \cdot X_2}$ فإن $r_{YX_1 \cdot X_2}$ له نفس إشارة $\hat{b}_1$ وكذلك $r_{YX_2 \cdot X_1}$ له نفس إشارة $\hat{b}_2$.

(د) تستخدم معاملات الارتباط الجزئية فى تحليل الانحدار المتعدد لتحديد الأهمية النسبية لكل متغير مفسر فى النموذج . والمتغير المستقل صاحب أعلى معامل ارتباط جزئى مع المتغير التابع يساهم أكثر من المتغيرات المستقلة الأخرى فى القدرة التفسيرية للنموذج ويدخل أولا فى تحليل الانحدار المتعدد خطوة – بخطوة . ولكن يجب ملاحظة أن معامل الارتباط الجزئى يعطى مقياساً لترتيب صافى الارتباط وليس مقياساً لقيمته ، فمجموع معاملات الارتباط الجزئى بين المتغير التابع وكل المتغيرات المستقلة لا يساوى 1 بالضرورة .

٧ - ٢٠ بالنسبة للانحدار المقدر فى المسألة ٧ - ٥ (أ) ، أوجد (أ) $r_{YX_1 \cdot X_2}$ ، (ب) $r_{YX_2 \cdot X_1}$. (جـ) هل تساهم X_1 أو X_2 أكثر فى القدرة التفسيرية للنموذج ؟

(أ) لإيجاد $r_{YX_1 \cdot X_2}$ فإننا نحتاج أولا إلى إيجاد r_{YX_1} ، r_{YX_2} و $r_{X_1X_2}$. باستخدام القيم من جدول ٧ - ٤ ، نحصل على

$$r_{YX_1} = \frac{\sum x_1 y}{\sqrt{\sum x_1^2}\sqrt{\sum y^2}} = \frac{-28}{\sqrt{60}\sqrt{40}} \cong -0.5715$$

$$r_{YX_2} = \frac{\sum x_2 y}{\sqrt{\sum x_2^2}\sqrt{\sum y^2}} = \frac{38}{\sqrt{74}\sqrt{40}} \cong 0.6984 \qquad (أ)$$

$$r_{X_1X_2} = \frac{\sum x_2 x_1}{\sqrt{\sum x_2^2}\sqrt{\sum x_1^2}} = \frac{-12}{\sqrt{74}\sqrt{60}} \cong -0.1801$$

وبالتالى

$$r_{YX_1 \cdot X_2} = \frac{r_{YX_1} - r_{YX_2}r_{X_1X_2}}{\sqrt{1-r_{X_1X_2}^2}\sqrt{1-r_{YX_2}^2}} = \frac{(-0.5715)-(0.6984)(-0.1801)}{\sqrt{1-(-0.1801)^2}\sqrt{1-0.6984^2}} \cong -0.6331$$

(ب) باستخدام قيم r_{YX_1} ، r_{YX_2} ، $r_{X_1X_2}$ السابق حسابها فى (أ) ، نحصل على

$$r_{YX_2 \cdot X_1} = \frac{r_{YX_2} - r_{YX_1}r_{X_1X_2}}{\sqrt{1-r_{X_1X_2}^2}\sqrt{1-r_{YX_1}^2}} = \frac{(0.6984)-(-0.5715)(-0.1801)}{\sqrt{1-(-0.1801)^2}\sqrt{1-(-0.5715)^2}} \cong 0.8072$$

(جـ) حيث أن $r_{YX_2 \cdot X_1}$ تتجاوز القيمة المطلقة للمعامل $r_{YX_1 \cdot X_2}$ فإننا نستنتج أن X_2 تساهم أكثر من X_1 فى القدرة التفسيرية للنموذج .

مسائل شاملة :

٧ - ٢١ يعطى جدول ٧ - ٦ الكمية المطلوبة (فرضاً) من سلعة ما ، Y وسعرها X_1 ، ودخل المستهلك ، X_2 من عام ١٩٧١ إلى ١٩٨٥ . (أ) هيىء انحدار OLS لهذه المشاهدات .

جدول ٧ - ٦ الكمية المطلوبة من سلعة ما ، سعرها ، ودخل المستهلك ، ١٩٧١ - ١٩٨٥

السنة	Y	X_1	X_2
1971	40	9	400
1972	45	8	500
1973	50	9	600
1974	55	8	700
1975	60	7	800
1976	70	6	900
1977	65	6	1,000
1978	65	8	1,100
1979	75	5	1,200
1980	75	5	1,300
1981	80	5	1,400
1982	100	3	1,500
1983	90	4	1,600
1984	95	3	1,700
1985	85	4	1,800

(ب) اختبر عند مستوى معنوية 5% المعنوية الإحصائية لمعالم الميل . (جـ) أوجد معامل الارتباط المتعدد غير المعدل والمعدل . (د) اختبر المعنوية الكلية للانحدار (هـ) أوجد معاملات الارتباط الجزئى وحدد أى متغير مستقل يساهم أكثر فى قدرة النموذج التفسيرية (و) أوجد معامل المرونة السعرية للطلب η_P ، والمرونة الدخلية للطلب ، η_M ، عند المتوسطات .
(ز) ضع جميع النتائج فى شكل ملخص مع تقريب كل الحسابات إلى 4 علامات عشرية .
(أ) يعطى جدول ٧ – ٧ الحسابات اللازمة لتوفيق الانحدار الخطى .

$$\hat{b}_1 = \frac{(\sum x_1 y)(\sum x_2^2) - (\sum x_2 y)(\sum x_1 x_2)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2} = \frac{(-505)(2{,}800{,}000) - (107{,}500)(-11{,}900)}{(60)(2{,}800{,}000) - (-11{,}900)^2} \cong 5.1061$$

$$\hat{b}_2 = \frac{(\sum x_2 y)(\sum x_1^2) - (\sum x_1 y)(\sum x_1 x_2)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2} = \frac{(107{,}500)(60) - (-505)(-11{,}900)}{(60)(2{,}800{,}000) - (-11{,}900)^2} \cong 0.1607$$

$$\hat{b}_0 = \bar{Y} - \hat{b}_1\bar{X}_1 - \hat{b}_2\bar{X}_2 = 70 - (-5.1061)(6) - (0.0167)(1{,}100) \cong 82.2666$$

$$\hat{Y} = 82.2666 - 5.1061\,X_1 + 0.0167\,X_2$$

(ب) ويمكننا إيجاد $\sum e_i^2$ بحساب R^2 أولا من جدول ٧ – ٧ :

$$R^2 = \frac{\hat{b}_1 \sum yx_1 + \hat{b}_2 \sum yx_2}{\sum y^2} = \frac{(-5.1061)(-505) + (0.0167)(107{,}500)}{4{,}600} \cong 0.9508$$

ولكن

$$R^2 = 1 - \frac{\sum e^2}{\sum y^2}$$

وبالتالى

$$\sum e^2 = (1 - R^2)\sum y^2 = (1 - 0.9508)4{,}600 \cong 226.32.$$

$$s_{\hat{b}_1}^2 = \frac{\sum e_i^2}{n-k}\,\frac{\sum x_2^2}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2}$$

$$= \frac{226.32}{15-3}\,\frac{2{,}800{,}000}{(60)(2{,}800{,}000) - (-11{,}900)^2} \cong 2.0011 \quad \text{and} \quad s_{\hat{b}_1} \cong 1.4146$$

$$s_{\hat{b}_2}^2 = \frac{\sum e_i^2}{n-k}\,\frac{\sum x_1}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2}$$

$$= \frac{226.32}{15-3}\,\frac{60}{(60)(2{,}800{,}000) - (-11{,}900)^2} \cong 0.00004 \quad \text{and} \quad s_{\hat{b}_2} \cong 0.0065$$

$$t_1 = \frac{\hat{b}_1}{s_{\hat{b}_1}} = \frac{-5.1061}{1.4146} \cong -3.6096 \quad \text{and} \quad t_2 = \frac{\hat{b}_2}{s_{\hat{b}_2}} = \frac{0.0167}{0.0065} \cong 2.5692$$

وبالتالى ، فإن كلا من b_1 و b_2 معنوية إحصائياً عند مستوى معنوية 5% .

(جـ) $R^2 = 0.9508$ (من (ب)) ، وعليه

$$\bar{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-k} = 1 - (1 - 0.9508)\frac{15-1}{15-3} \cong 0.9426$$

(د)

$$F_{k-1,\,n-k} = \frac{R^2/(k-1)}{(1-R^2)/(n-k)} = \frac{0.9508/(3-1)}{(1-0.9508)/(15-3)} \cong 115.9512$$

وبالتالى ، فإن R^2 تختلف معنويا عن الصفر عند مستوى معنوية 5% .

جدول ٧ – ٧ انحدار كمية الطلب : الحسابات

السنة	Y	X_1	X_2	y	x_1	x_2	yx_1	yx_2	x_1x_2	x_1^2	x_2^2	y^2
1971	40	9	400	−30	3	−700	−90	21,000	−2,100	9	490,000	900
1972	45	8	500	−25	2	−600	−50	15,000	−1,200	4	360,000	625
1973	50	9	600	−20	3	−500	−60	10,000	−1,500	9	250,000	400
1974	55	8	700	−15	2	−400	−30	6,000	−800	4	160,000	225
1975	60	7	800	−10	1	−300	−10	3,000	−300	1	90,000	100
1976	70	6	900	0	0	−200	0	0	0	0	40,000	0
1977	65	6	1,000	−5	0	−100	0	500	0	9	10,000	25
1978	65	8	1,100	−5	2	0	−10	0	0	4	0	25
1979	75	5	1,200	5	−1	100	−5	500	−100	1	10,000	25
1980	75	5	1,300	5	−1	200	−5	1,000	−200	1	40,000	25
1981	80	5	1,400	10	−1	300	−10	3,000	−300	1	90,000	100
1982	100	3	1,500	30	−3	400	−90	12,000	−1,200	9	160,000	900
1983	90	4	1,600	20	−2	500	−40	10,000	−1,000	4	250,000	400
1984	95	3	1,700	25	−3	600	−75	15,000	−1,800	9	360,000	625
1985	85	4	1,800	15	−2	700	−30	10,500	−1,400	4	490,000	225
$n = 15$	$\sum Y = 1,050$ $\bar{Y} = 70$	$\sum X_1 = 90$ $\bar{X}_1 = 6$	$\sum X_2 = 16,500$ $\bar{X}_2 = 1,100$	$\sum y = 0$	$\sum x_1 = 0$	$\sum x_2 = 0$	$\sum yx_1 = -505$	$\sum yx_2 = 107,500$	$\sum x_1x_2 = -11,900$	$\sum x_1^2 = 60$	$\sum x_2^2 = 2,800,000$	$\sum y^2 = 4,600$

(هـ) لإيجاد $r_{YX_1 \cdot X_2}$ و $r_{YX_2 \cdot X_1}$ ، فيجب أولا إيجاد (من جدزول ٧ – ٧) .

$$r_{YX_1} = \frac{\sum x_1 y}{\sqrt{\sum x_1^2}\sqrt{\sum y^2}} = \frac{-505}{\sqrt{60}\sqrt{4,600}} \cong 0.9613$$

$$r_{YX_2} = \frac{\sum x_2 y}{\sqrt{\sum x_2^2}\sqrt{\sum y^2}} = \frac{107,500}{\sqrt{2,800,000}\sqrt{4,600}} \cong 0.9472$$

$$r_{X_1X_2} = \frac{\sum x_2 x_1}{\sqrt{\sum x_2^2}\sqrt{\sum x_1^2}} = \frac{-11,900}{\sqrt{2,800,000}\sqrt{60}} \cong 0.9181$$

$$r_{YX_1 \cdot X_2} = \frac{r_{YX_1} - r_{YX_2} r_{X_1X_2}}{\sqrt{1 - r_{X_1X_2}^2}\sqrt{1 - r_{YX_2}^2}} = \frac{(-0.9613) - (0.9472)(-0.9181)}{\sqrt{1 - (-0.9181)^2}\sqrt{1 - (-0.9472)^2}} \cong -0.7213$$

$$r_{YX_2 \cdot X_1} = \frac{r_{YX_2} - r_{YX_1} r_{X_1X_2}}{\sqrt{1 - r_{X_1X_2}^2}\sqrt{1 - r_{YX_1}^2}} = \frac{(0.9472) - (-0.9613)(-0.9181)}{\sqrt{1 - (-0.9181)^2}\sqrt{1 - (-0.9613)^2}} \cong 0.5919$$

وبالتالى فإن X_1 تساهم أكثر من X_2 فى القدرة التفسيرية للنموذج

(و)

$$\eta_P = \hat{b}_1 \frac{\bar{X}_1}{\bar{Y}} = -5.1061 \frac{6}{70} \cong 0.4377$$

$$\eta_M = \hat{b}_2 \frac{\bar{X}_2}{\bar{Y}} = 0.0167 \frac{1,100}{70} \cong 0.2624$$

(ز)

$$\hat{Y}_1 = 82.2666 - 5.1061 X_1 + 0.0167 X_2 \qquad R^2 = 0.9508 \quad \bar{R}^2 = 0.9426 \quad F_{2,12} = 155.9512$$

t values (−3.6096) (2.5692)

$$r_{YX_1 \cdot X_2} = 0.7023 \quad r_{YX_2 \cdot X_1} = 0.8434$$

$$\eta_P = -0.4377 \quad \eta_M = 0.2624$$

٧ – ٢٢ جدول ٧ – ٨ هو نفس جدول ٧–٦ فيما عدا أنه يشمل سعر سلعة بديلة بالدولار ، X_3 كمتغير مستقل ثالث . جدول ٧ – ٩ يعطى صورة من مخرجات الكمبيوتر للانحدار الخطى المتغير Y على X_1 ، X_2 ، X_3 باستخدام Statistical Package for the Social Sciences) SPSS أكثر برامج الكمبيوتر شيوعاً) .

جدول ٧ – ٨ الكمية المطلوبة ، السعر ، دخل المستهلك ، وسعر سلعة بديلة

السنة	Y	X_1	X_2	X_3
1971	40	9	400	10
1972	45	8	500	14
1973	50	9	600	12
1974	55	8	700	13
1975	60	7	800	11
1976	70	6	900	15
1977	65	6	1,000	16
1978	65	8	1,100	17
1979	75	5	1,200	22
1980	75	5	1,300	19
1981	80	5	1,400	20
1982	100	3	1,500	23
1983	90	4	1,600	18
1984	95	3	1,700	24
1985	85	4	1,800	21

أجب عن الأسئلة التالية من خلال فحص مخرجات الكمبيوتر فى جدول ٧ – ٩ . (أ) اكتب معادلة انحدار OLS مع قيم t ، R^2 ، $\bar{R}^2$ نسبة F مع درجات الحرية ، الخطأ المعيارى للانحدار ، ومجموع مربعات البواقى ، وفسر النتائج .
(ب) كيف أجرى برنامج SPSS ؟

(أ) من صفحة ٤ من مخرجات الكمبيوتر فى جدول ٧ – ٩ ، نحصل على :

$$t_1 = \frac{\hat{b}_1}{s_{\hat{b}_1}} = \frac{-4.9281}{1.6111} = -3.059 \qquad t_2 = \frac{\hat{b}_2}{s_{\hat{b}_2}} = \frac{0.0159}{0.0074} = 2.149 \qquad t_3 = \frac{\hat{b}_3}{s_{\hat{b}_3}} = \frac{0.1748}{0.6367} = 0.275$$

وعليه

$$Y = 79.1063 - 4.9281X_1 + 0.0159X_2 + 0.1748X_3 \qquad R^2 = 0.95 \quad \bar{R}^2 = 0.94 \quad F_{3,11} = 71.13$$
$$(-3.059) \qquad (2.149) \qquad (0.275) \qquad s = 4.53 \quad SEE = 225.49$$

إشارات $\hat{b}_1$ ، $\hat{b}_2$ ، $\hat{b}_3$ تتفق مع ما هو متوقع طبقاً للنظرية الإقتصادية . ولكن $\hat{b}_1$ فقط معنوية إحصائياً عند مستوى معنوية 5% ، بينما كادت $\hat{b}_2$ أن تكون معنوية . من نسبة F ، ننتهى إلى أن R^2 (والانحدار ككل) معنوى إحصائياً عند مستوى 5% . يجب ملاحظة أنه فى صفحة ٤ من جدول ٧ – ٩ ، R المتعددة هى الجذر التربيعى للمقدار R تربيع وبالتالى تشير إلى معامل الارتباط المتعدد r .

جدول ٧ – ٩ مخرجات الكمبيوتر SPSS لانحدار الطلب

```
SPSS Batch System                                            11-Dec-80     Page  1

SPSS for DECSYSTEM-20, Version H, Release 8.0 (17-Dec-79)

Default SPACE allocation!      Allows for!        98 Transformations
WORKSPACE      17920 words                       394 RECODE values + LAG variables
TRANSPACE       2560 words                      1576 IF/COMPUTE operations

             1 RUN NAME        MULTIPLE REGRESSION
             2 FILE NAME       MULTRG
             3 VARIABLE LIST   N,Y,X1,X2,X3
             4 INPUT FORMAT    FREEFIELD
             5 N OF CASES      15
             6 REGRESSION      VARIABLES=Y,X1,X2,X3/
             7                 REGRESSION=Y WITH X1 TO X3 (2) RESID=0/
             8 OPTION          11,12
             9 STATISTICS      1 TO 8

 ***** REGRESSION problem requires    128 words WORKSPACE, not including residuals *****

            10 READ INPUT DATA

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

MULTIPLE REGRESSION                                          11-Dec-80     Page  2

File   MULTRG   (Creation date = 11-Dec-80)

Variable          Mean       Standard dev      Cases

Y              70.0000         18.1265           15
X1              6.0000          2.0702           15
X2           1100.0000        447.2136           15
X3             17.0000          4.4721           15

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

MULTIPLE REGRESSION                                          11-Dec-80     Page  3

File   MULTRG   (Creation date = 11-Dec-80)

Correlation coefficients

A value of 99.00000 is printed
if a coefficient cannot be computed.


              Y          X1          X2          X3

Y          1.00000    -0.96125     0.94722     0.88995
X1        -0.96125     1.00000    -0.91810    -0.88724
X2         0.94722    -0.91810     1.00000     0.88571
X3         0.88995    -0.88724     0.88571     1.00000

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
```

تابع جدول ٧ — ٩

MULTIPLE REGRESSION — 11-Dec-80 — Page 4

File MULTRG (Creation date = 11-Dec-80)

* M U L T I P L E R E G R E S S I O N * * * * * * * * * * * * * * Variable list 1 / Regression list 1

Dependent variable: Y

Variable(s) entered on step number 1: X3, X1, X2

			Analysis of variance	DF	Sum of squares	Mean square	F
Multiple R	0.97518		Regression	3.	4374.50821	1458.16940	71.13280
R square	0.95098		Residual	11.	225.49179	20.49925	
Adjusted R square	0.93761						
Standard error	4.52761						

------------------ Variables in the equation ------------------

Variable	B	Beta	Std error B	F
X3	0.1747977D+00	0.04313	0.63672	0.075
X1	-0.4928058D+01	-0.56282	1.61108	9.357
X2	0.1590040D-01	0.39229	0.00741	4.604
(Constant)	0.7910634D+02			

------------- Variables not in the equation -------------

Variable	Beta in	Partial	Tolerance	F

All variables are in the equation

Statistics which cannot be computed are printed as all nines.

- -

MULTIPLE REGRESSION — 11-Dec-80 — Page 5

File MULTRG (Creation date = 11-Dec-80)

* M U L T I P L E R E G R E S S I O N * * * * * * * * * * * * * * Variable list 1 / Regression list 1

Dependent variable: Y

Summary table

Variable	Multiple R	R square	Rsq change	Simple R	B	Beta
X3	0.88995	0.79200	0.79200	0.88995	0.1747977D+00	0.04313
X1	0.96461	0.93047	0.13846	-0.96125	-0.4928058D+01	-0.56282
X2	0.97518	0.95098	0.02051	0.94722	0.1590040D-01	0.39229
(Constant)					0.7910634D+02	

- -

MULTIPLE REGRESSION — 11-Dec-80 — Page 6

* * * * * Note change in formula for Standardized Residuals as of 17 Dec 79 * * * * *
It was (Residual/Std. Dev. of Dep. Variable)
It is now (Residual/Std. Error of Regression)

* * * * * REGRESSION problem requires 1094 words WORKSPACE including residuals * * * * *

- -

MULTIPLE REGRESSION — 11-Dec-80 — Page 7

File MULTRG (Creation date = 11-Dec-80)

* M U L T I P L E R E G R E S S I O N *

Dependent variable: Y from variable list 1, regression list 1

Plot of standardized residual: -2.0 -1.0 0.0 1.0 2.0

SEQNUM	Observed Y	Predicted Y	Residual
1	50.00000	46.39164	3.608363
2	95.00000	95.54800	-0.5479991
3	75.00000	78.45773	-3.457734
4	65.00000	60.14388	4.856115
5	100.0000	92.19312	7.806880
6	85.00000	91.68559	-6.685589
7	70.00000	66.47032	3.529676
8	40.00000	42.86196	-2.861960
9	45.00000	50.07925	-5.079249
10	90.00000	87.98112	2.018885
11	55.00000	53.08453	1.915468
12	80.00000	80.22257	-0.2225719
13	75.00000	77.39209	-2.392086
14	65.00000	68.23516	-3.235162
15	60.00000	59.25304	0.7469649

Durbin-Watson test of residual differences compared by case order (SEQNUM).

Variable list 1, regression list 1. Durbin-Watson test 2.39460

- -

وتشير بيتا إلى المعاملات المعيارية ، أو المعاملات المقدرة مضروبة فى نسبة الانحراف المعيارى لمتغير مستقل معين إلى الانحراف المعيارى للمتغير التابع . قيمة F لكل معامل ليست إلا مربع قيمة t لكل معامل .

(ب) الخطوة الأولى لإجراء برنامج كمبيوتر هى « الدخول » من إحدى محطات الكمبيوتر ويحتاج هذا إلى رقم حساب « وكلمة مرور » (والتى تحصل عليها من مركز الكمبيوتر فى جامعتك ، أو شركتك) . لاستخدام SPSS فإنك تستخدم أمراً ، يمدك به أيضاً مركز الكمبيوتر . الخطوات التالية موجودة فى صفحة ١ (السطور من ٥ إلى ١٥) من مخرجات الكمبيوتر فى جدول ٧ - ٩ . وهذه فى معظمها لا تحتاج إلى شرح . فى السطر الرابع (INPUT FORMAT) تشير FREEFIELD إلى الطريقة التى يغذى بها الكمبيوتر بالبيانات ، حيث تدخل قيمة كل متغير فى التسلسل المشار إليه فى سطر ٣ (أى .50 9 600 12 95 3 1, 700 24 75 etc) فى سطر ٧ ، تشير الأقواس إلى أن المطلوب هو انحدار متعدد عادى ، بينما تطلب RESID = 0 حساب البواقى المعيارية وقيم التنبؤ للمتغير Y . ويطلب OPITON 11, 12 على سطر ٨ كتابة البواقى المعيارية وقيم التنبؤ المتغير Y (المعطاة فى صفحة ٧ من مخرجات الكمبيوتر) . وتطلب STATISTICS 1 to 8 فى سطر ٩ كتابة المتوسط ، الانحراف المعيارى وعدد الحالات الصحيحة (السنوات) لكل متغير (المعطاة فى صفحة ٢ من مخرجات الكمبيوتر ؛ مصفوفة الارتباط البسيط (المعطاة فى صفحة ٣) ؛ كل الإحصائيات فى صفحة ٤ (السابق مناقشتها فى (أ)) ، وإحصائية ديربين - واتسون . ارجع إلى دليل SPSS فى مركز الكمبيوتر . وأخيراً ، فإن الجدول الموجز فى صفحة ٥ (ويتضمن كل مخرجات الكمبيوتر بشكل أوتوماتيكى) يعالج المتغيرات كما لو كانت أدخلت الكمبيوتر واحداً وراء الآخر وليس معاً .

مسائل إضافية

النموذج الخطى ذو المتغيرات الثلاثة :

٧ - ٢٣ جدول ٧ - ١٠ امتداد لجدول ٦ - ١٢ ويعطى مشاهدات عن Y ، X_1 ، X_2 . أوجد معادلة انحدار OLS للمتغير Y على X_1 و X_2 .

الإجابة : $\hat{Y}_i = 4.76 + 5.29X_{1i} + 2.13X_{2i}$

جدول ٧ - ١٠ مشاهدات عن Y ، X_1 ، X_2

n	1	2	3	4	5	6	7	8	9	10
Y	20	28	40	45	37	52	54	43	65	56
X_1	2	3	5	4	3	5	7	6	7	8
X_2	5	6	6	5	5	7	6	6	7	7

٧ - ٢٤ بالرجوع إلى معادلة انحدار OLS المقدرة فى المسألة ٧ - ٢٣ فسر (أ) $\hat{b}_0$ ، (ب) $\hat{b}_1$ ، (جـ) $\hat{b}_2$.

الإجابة : (أ) $\hat{b}_0 = 4.76$ هى الثابت أو مقطع Y ، $\hat{Y}_i = \hat{b}_0 = 4.76$. ، عند $X_{1i} = X_{2i} = 0$

(ب) $b_1 = 5.29$ تشير إلى أن زيادة قدرها الوحدة فى X_1 (مع تثبيت X_2) ينتج عنها زيادة فى $\hat{Y}_i$ قدرها 5.29 وحدة

(جـ) $b_2 = 2.13$ ، تشير إلى أن زيادة قدرها الوحدة فى X_2 (مع تثبيت X_1) ينتج عنها زيادة فى $\hat{Y}_i$ قدرها 2.13 وحدة .

اختبارات معنوية تقديرات المعالم :

٧ - ٢٥ بالرجوع إلى بيانات جدول ٧ - ١٠ ، أوجد (أ) S^2 ، (ب) $s^2_{\hat{b}_1}$ و $s_{\hat{b}_1}$ ، (جـ) $s^2_{\hat{b}_2}$ و $s_{\hat{b}_2}$.

الإجابة : (أ) $S^2 = 50$ (ب) $s^2_{\hat{b}_1} \cong 3.16$ and $s_{\hat{b}_1} \cong 1.78$ (جـ) $s^2_{\hat{b}_2} \cong 18.95$ و $s_{\hat{b}_2} \cong 4.35$

٧ – ٢٦ اختبر عند مستوى معنوية 5% كلا من (أ) b_1 ، (ب) b_2 فى المسألة ٧ – ٢٣ .

الإجابة : (أ) b_1 معنوية إحصائياً عند مستوى 5% .

(ب) b_2 ليست معنوية إحصائياً عند مستوى 5% .

٧ – ٢٧ كون فترة الثقة 95% لكل من (أ) b_1 ، (ب) b_2 فى المسألة ٧ – ٢٣ .

الإجابة : (أ) $1.08 \le b_1 \le 9.50$ (ب) $-8.16 \le b_2 \le 12.42$

معامل التحديد المتعدد :

٧ – ٢٨ بالنسبة لانحدار OLS المقدرة فى المسألة ٧ – ٢٣ ، أوجد (أ) R^2 ، (ب) $\bar{R}^2$ (ج) هل يجب أن تدخل X_2 فى الانحدار ؟

الإجابة : (أ) $R^2 \cong 0.79$ (باستخدام $[R^2 = 1 - (\sum e_i^2/\sum y_i^2)]$ (ب) $\bar{R}^2 \cong 0.73$ (ج) حيث أن b_2 وجدت غير معنوية إحصائياً (فى المسألة ٧ – ٢٦ (ب)) ، R^2 نقصت من $R^2 = \bar{R}^2 = 0.77$ عندما كانت X_1 المتغير المستقل الوحيد (انظر المسألة ٦ – ٤١ (أ)) إلى $\bar{R}^2 = 0.73$ (أعلاه) ، فإن X_2 يجب ألا تدخل الانحدار .

٧ – ٢٩ بالنسبة إلى $R^2 = 0.60$ ، $n = 10$ ، $K = 1$ ، أوجد $\bar{R}^2$.

الإجابة : $\bar{R}^2 = 0.60$

٧ – ٣٠ بالنسبة إلى $R^2 = 0.60$ ، $n = 10$ ، $K = 2$ ، أوجد $\bar{R}^2$.

الإجابة : $\bar{R}^2 = 0.55$

٧ – ٣١ بالنسبة إلى $R^2 = 0.60$ ، $K = 2$ (كما فى المسألة ٧ – ٣٠) ولكن $n = 100$ ، أوجد $\bar{R}^2$

الإجابة : $\bar{R}^2 = 0.596$

٧ – ٣٢ بالنسبة إلى $R^2 = 0.40$ ، $n = 10$ ، $K = 5$ ، أوجد $\bar{R}^2$.

الإجابة : $\bar{R}^2 = -0.08$ (ولكنها تفسر على أنها تساوى الصفر) .

اختبار المعنوية الكلية للانحدار :

٧ – ٣٣ من انحدار OLS المقدر فى المسألة ٧ – ٢٣ ، أوجد (أ) التباين المفسر ، (ب) التباين غير المفسر أو تباين البواقى (ج) نسبة أو إحصائية F .

الإجابة : (أ) $\sum \hat{y}^2/(k-1) \cong 649$ (ب) $\sum e^2/(n-k) = 50$ (ج) $F_{2,7} = 12.98$

٧ – ٣٤ اختبر المعنوية الإجمالية لانحدار OLS المقدر فى المسألة ٧ – ٢٣ عند (أ) مستوى 5% (ب) مستوى 1% .

الإجابة : (أ) حيث أن نسبة F المحسوبة (12.98) تتجاوز قيمة F النظرية أو الجدولية (4.74) عند $\alpha = 0$ ودرجات حرية 2 ، 7 ، فإننا نقبل الفرض بأن معالم انحدار OLS المقدرة هى معاً معنوية عند مستوى 5% .

(ب) حيث أن قيمة F الجدولية عند مستوى $\alpha = 0.01$ هى $F = 9.55$ ، يقبل الفرض البديل عند مستوى معنوية 1% أيضاً .

معاملات الارتباط الجزئية :

٧ – ٣٥ بالنسبة لانحدار OLS المقدر فى المسألة ٧ – ٢٣ ، أوجد (أ) $r_{YX_1 \cdot X_2}$ ، (ب) $r_{YX_2 \cdot X_1}$ (ج) أى متغير مستقل يساهم أكثر فى القدرة التفسيرية للنموذج ؟

الإجابة : (أ) $= 0.74$ (ب) $r_{YX_2 \cdot X_1} = 0.13$ (ج) X_1

مسألة شاملة :

٧ — ٣٦ جدول ٧ — ١١ امتداد لجدول ٦ — ١٣ ويعطى بيانات عينة عشوائية من 12 من الأسر عن عدد الأطفال فى الأسرة Y ، وعدد الأطفال الذين قالوا إنهم كانوا يرغبون فى إنجابهم وقت الزواج ، X_1 ، وعدد سنوات تعليم الزوجة ، X_2 . (أ) أوجد معادلة انحدار OLS للمتغير Y على X_1 و X_2 . (ب) احسب قيم t واختبر عند مستوى 5% المعنوية الإحصائية لمعالم الميل . (جـ) أوجد معامل الارتباط المتعدد غير المعدل والمعدل (د) اختبر المعنوية الإجمالية للانحدار (هـ) أوجد معاملات الارتباط الجزئية وحدد أى المتغيرات المستقلة يساهم أكثر فى القدرة التفسيرية للنموذج . قم بجميع الحسابات إلى رقمين عشريين .

جدول ٧ — ١١ عدد الأطفال فى الأسرة وعدد الأطفال المرغوب فيهم وتعليم الزوجة

الأسرة	1	2	3	4	5	6	7	8	9	10	11	12
Y	4	3	0	4	4	3	0	4	3	1	3	1
X_1	3	3	0	2	2	3	0	3	2	1	3	2
X_2	12	14	18	10	10	14	18	12	15	16	14	15

الإجابة : (أ) $\hat{Y} = 6.90 + 0.53X_1 - 0.39X_2$ (ب) حيث أن $t_1 = 3.12$ و $t_2 = -5.57$ ، فإن كلا من $\hat{b}_1$ و $\hat{b}_2$ معنوية إحصائياً عند مستوى 5% . (جـ) $R^2 = 0.92$ و $\bar{R}^2 = 0.90$ (د) حيث أن $F_{2,9} = 51.31$ فإن R^2 معنوية إحصائياً عند مستوى 5% (هـ) $r_{YX_2 \cdot X_1} = 0.87$ و $r_{YX_1 \cdot X_2} = 0.71$ وبالتالى فإن X_2 تساهم أكثر من X_1 فى القدرة التفسيرية للنموذج .

الفصل الثامن

اساليب وتطبيقات اخرى فى تحليل الانحدار

٨ــ١ شكل الدالة

كثيراً ما توحى النظرية أو شكل الانتشار بوجود علاقة غير خطية . ومن الممكن تحويل بعض الدوال غير الخطية إلى دوال خطية حتى يمكن تطبيق طريقة OLS . ويوضح جدول (٨ – ١) بعضاً من أكثر هذه الدوال شيوعاً وتحويلاتها . وتطبيق طريقة OLS على العلاقات الخطية المحولة يعطى تقديرات غير متحيزة للميل . فى معادلة (٨ – ١) ، b_1 هى مرونة Y بالنسبة إلى X

جدول (٨ – ١) أشكال الدالة وتحويلاتها

المعادلة	الشكل	التحويلية	الدالة
(٨ – ١)	لوغاريتمى مزدوج	$Y^* = b_0^* + b_1 X^* + u$	$Y = b_0 X^{b_1} e^u$
(٨ – ٢)	نصف لوغاريتمى	$Y^* = b_0 + b_1 X + u$	$\ln Y = b_0 + b_1 X + u$
(٨ – ٣)	مقلوب	$Y = b_0 + b_1 Z + u$	$Y = b_0 + \frac{b_1}{X} + u$
(٨ – ٤)	تربيعى	$Y = b_0 + b_1 X + b_2 W + u$	$Y = b_0 + b_1 X + b_2 X^2 + u$

حيث $Y^* = \ln Y, \quad b_0^* = \ln b_0, \quad X^* = \ln X, \quad u = \ln e^u, \quad Z = 1/X, \quad W = X^2$

$\ln$ = اللوغاريتم الطبيعى الأساسى $e \cong 2.718$ = ln

مثال (١) : افترض أننا سلمنا مقدماً بمعادلة طلب من الشكل

$$Y = b_0 X_1^{b_1} X_2^{b_2} e^u$$

حيث Y = الكمية المطلوبة من السلعة

X_1 = سعرها

X_2 = دخول المستهلكين

باستخدام بيانات جدول (٧ – ٦) وتطبيق طريقة OLS لدالة الاستهلاك هذه بعد تحويلها إلى شكل خطى لوغاريتمى مزدوج ، نحصل على :

$$\ln Y = 1.96 - 0.26 \ln X_1 + 0.39 \ln X_2 \qquad R^2 = 0.97$$
$$(-3.54) \qquad (6.64)$$

حيث 0.26 — و 0.39 هما ، على الترتيب ، تقديرات غير متحيزة للمرونة السعرية والدخلية للطلب (انظر المسألة ٨ – ٢) ويبدو أن التوفيق هنا أفضل منه فى الحالة الخطية (انظر المسألة (٧ – ٢٠ (ز)) .

٨ــ٢ المتغيرات الصورية

يمكن تقديم المتغيرات المفسرة الكيفية (مثل الحرب مقابل السلام ، فترات الإضراب مقابل فترات عدم الإضراب ، الذكور مقابل الإناث ، الخ) فى تحليل الانحدار بتعيين قيمة 1 لأحد التصنيفين (الحرب مثلا) و 0 للتصنيف الآخر (السلام ، مثلا) . وتسمى هذه

بالمتغيرات الصورية وتعامل معاملة المتغيرات الأخرى . ويمكن استخدام المتغيرات الصورية للامساك بالمتغيرات (التنقلات) فى ثابت المعادلة (معادلة(٨ — ٥)) ، والمتغيرات فى الميل (معادلة (٨ — ٦)) ، أو التغيرات فى كليهما (معادلة (٨ — ٧)) :

$$Y = b_0 + b_1X + b_2D + u \qquad (٨-٥)$$

$$Y = b_0 + b_1X + b_2XD + u \qquad (٨-٦)$$

$$Y = b_0 + b_1X + b_2D + b_3XD + u \qquad (٨-٧)$$

حيث D هى 1 لأحد التصنيفين و 0 للتصنيف الآخر وX هى المتغير المفسر الكمى المعتاد . ويمكن استخدام المتغيرات الصورية للإمساك بالفروق بين أكثر من تصنيفين ، مثل المواسم والمناطق (معادلة (٨ — ٨)) :

$$Y = b_0 + b_1X + b_2D_1 + b_3D_2 + b_4D_3 + u \qquad (٨-٨)$$

حيث b_0 هى المقطع للموسم أو الإقليم الأول و D_1 ، D_2 وD_3 تشير على الترتيب إلى المواسم أو الأقاليم 2 ، 3 و4 . لاحظ أنه بالنسبة لأى عدد من التصنيفات k ، فإننا نحتاج إلى $k-1$ من المتغيرات الصورية (انظر المسائل ٨ — ٩ ، ٨ — ٢٦ ، و ٨ — ٢٧) . بالنسبة للمتغيرات التابعة الكيفية ، انظر المسألة (٨ — ١٠) .

مثال (٢) : يعطى جدول (٨ — ٢) إجمالى الاستثمار الخاص المحلى ، Y ، والناتج القومى الإجمالى ، X ، بالبليون دولار وبالأسعار الجارية ، فى الولايات المتحدة من ١٩٣٩ إلى ١٩٥٤ ، باستخدام $D = 1$ لسنوات الحرب (١٩٤٢ — ١٩٤٥) و $D = 0$ لسنوات السلام ، نحصل على

$$\hat{Y} = -2.58 - \underset{(10.79)}{0.16X} - \underset{(-6.82)}{20.81D} \qquad R^2 = 0.94$$

D معنوية إحصائياً عند مستوى 5% . أى أن $\hat{b}_0 = -2.58$ لوقت السلم و 23.39 — لوقت الحرب ، بينما $\hat{b}_1 = 0.16$ هى معامل الميل المشترك . (لاختبارات الاختلاف فى الميل ، وكذلك الاختلاف فى المقطع وفى الميل ، (أنظر المسائل (٨ — ٧) ، (٨ — ٨) ، (٨ — ٢٤) ، و (٨ — ٢٥) .

جدول (٨ — ٢) إجمالى الاستثمار الخاص المحلى والناتج القومى الإجمالى (ببلايين الدولارات)
الولايات المتحدة (١٩٣٩ — ١٩٥٤)

السنة	1939	1940	1941	1942	1943	1944	1945	1946	1947	1948	1949	1950	1951	1952	1953	1954
Y	9.3	13.1	17.9	9.9	5.8	7.2	10.6	30.7	34.0	45.9	35.3	53.8	59.2	52.1	53.3	52.7
X	90.8	100.0	124.9	158.3	192.0	210.5	212.3	209.3	232.8	259.1	258.0	286.2	330.2	347.2	366.1	366.3

المصدر : التقرير الاقتصادى للرئيس ، مكتب حكومة الولايات المتحدة للطباعة ، واشنطن ١٩٨٠ ، صفحة ٢٠٣

٨—٣ نماذج فترات الابطاء الموزعة

غالباً ما تكون قيمة المتغير التابع الحالية دالة فى أو تعتمد على مجموع مرجح للقيم الحالية t ؟ والماضية للمتغير المستقل (وحد الخطأ) ، مع تعيين أوزان مختلفة عادة للفترات الزمنية المختلفة :

$$Y_t = a + b_0X_t + b_1X_{t-1} + b_2X_{t-2} + \cdots + u_t \qquad (٨-٩)$$

وتقدير نموذج فترات الإبطاء الموزعة (معادلة (٨ — ٩)) يمثل صعوبتين . الأولى ، أن بيانات مشاهدة أو فترة زمنية تضيع لكل قيمة مبطأة للمتغير X والثانية ، إن قيم المتغيرات المستقلة X على الأرجح سوف تكون مرتبطة بعضها ببعض وبالتالى سوف يصعب عزل تأثير كل X على Y

ويمكن التخلص من هذه الصعوبات بأن نشتق من معادلة (٨ – ٩) نموذج إبطاء كويك ، (معادلة (٨ – ١٠)) ، والذى يفترض أن الأوزان تتناقص كمتوالية هندسية (أنظر المسألة ٨ – ١٢) :

$$Y_t = a(1-\lambda) + b_0 X_t + \lambda Y_{t-1} + v_t \qquad (١٠-٨)$$

حيث $0 < \lambda < 1$ و $v_t = u_t - \lambda u_{t-1}$.

ولكن معادلة (٨ – ١٠) تخرق اثنين من فروض نموذج OLS وتؤدى إلى مقدرات متحيزة وغير متسقة ومن ثم تحتاج إلى تعديل (أنظر قسم ٩ – ٣) .

وكبديل ، يمكن استخدام نموذج إبطاء ألمون . ويسمح هذا بهيكل إبطاء أكثر مرونة ، ويمكن تقريبه عملياً باستخدام كثيرة حدود تزيد درجتها عن عدد نقاط التحول فى الدالة بواحد على الأقل (انظر مسألة ٨ – ١٣) . وبافتراض إبطاء لثلاث فترات (معادلة ٨ – ١١) على شكل معادلة تربيعية (معادلة ٨ – ١٢) ، يمكننا اشتقاق معادلة (٨ – ١٣) (انظر المسألة ٨ – ١٥) :

$$Y_t = a + b_0 X_t + b_1 X_{t-1} + b_2 X_{t-2} + b_3 X_{t-3} + u_t \qquad (١١-٨)$$

حيث

$$b_i = c_0 + c_1 i + c_2 i^2 \qquad (١٢-٨)$$

بحيث أن

$$Y_t = a + c_0 Z_{1t} + c_1 Z_{2t} + c_2 Z_{3t} + v_t \qquad (١٣-٨)$$

حيث

$$Z_{1t} = \sum_{i=0}^{3} X_{t-i} \qquad Z_{2t} = \sum_{i=1}^{3} i X_{t-i} \qquad \text{و} \qquad Z_{3t} = \sum_{i=1}^{3} i^2 X_{t-i}$$

ونحصل على قيم المعاملات $\hat{b}_i$'s فى معادلة (٨ – ١١) بالتعويض بالقيم المقدرة للمعاملات c_0 ، c_1 و c_3 من معادلة (٨ – ١٣) فى معادلة (٨ – ١٢) (انظر المسألة ٨ – ١٦) .

مثال (٣) : يعطى جدول (٨ – ٣) مستوى الواردات Y ، والدخل القومى الإجمالى X ، كليهما ببلايين الدولارات ، للولايات المتحدة للسنوات من ١٩٦٠ إلى ١٩٧٩ . بتوفيق نموذج كويك ، نحصل على

$$\hat{Y}_t = -27.45 + \underset{(2.52)}{0.06}\, X_t + \underset{(2.66)}{0.60}\, Y_{t-1} \qquad R^2 = 0.99$$

حيث $\hat{\lambda} = 0.60$ و $\hat{a}\,(1 - 0.60) = -27.45$ أى أن $\hat{a} = -68.63$.

جدول (٨ – ٣) الواردات والدخل القومى الإجمالى (ببلايين الدولارات) : الولايات المتحدة ، ١٩٦٠ – ١٩٧٩

Year	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969
Imports	23.2	23.1	25.2	26.4	28.4	32.0	37.7	40.6	47.7	52.9
GNP	506.0	523.3	563.8	594.7	635.7	688.1	753.0	796.3	868.5	935.5
Year	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
Imports	58.5	64.0	75.9	94.4	131.9	126.9	155.4	185.8	217.5	260.9
GNP	982.4	1,063.4	1,171.1	1,306.6	1,412.9	1,528.8	1,702.2	1,899.5	2,127.6	2,368.5

المصدر : التقرير الاقتصادي للرئيس . مكتب حكومة الولايات المتحدة للطباعة . واشنطن ١٩٨٠ صفحة ٢٠٣

٨-٤ التنبـــؤ

يشير التنبؤ إلى تقدير قيمة المتغير التابع ، Y_F ، بمعلومية القيمة الفعلية أو المتوقعة للمتغير المستقل ، X_F . ويمثل تباين خطأ التنبؤ ، σ_F^2 بالآتى :

$$\sigma_F^2 = \sigma_u^2\left[1 + \frac{1}{n} + \frac{(X_F - \bar{X})^2}{\sum(X_i - \bar{X})^2}\right] \qquad (٨-١٤)$$

حيث n هى عدد المشاهدات و σ_u^2 هى تباين u . وحيث أن σ_u^2 تكون عادة غير معلومة ، فإننا نستخدم s^2 كتقدير غير متحيز للتباين σ_F^2 ، فيكون تباين خطأ التنبؤ

$$s_F^2 = s^2\left[1 + \frac{1}{n} + \frac{(X_F - \bar{X})^2}{\sum(X_i - \bar{X})^2}\right] \qquad (٨-١٥)$$

وتكون فترة الثقة 95% للتنبؤ ، Y_F ، هى

$$\hat{Y}_F \pm t_{0.025} s_F$$

حيث $Y_F = \hat{b}_0 + \hat{b}_1 X_F$ وتشير t إلى توزيع t بدرجات حرية $n - 2$.

مثال (٤) : بالعودة إلى مثال الحنطة – السماد فى الفصل السادس ، نتذكر أن $\hat{Y}_i = 27.12 + 1.66 X_i$ ، $n = 10$ ، $\bar{X} = 18$ ، $\Sigma(X_i - \bar{X})^2 = 576$ ، (من مثال (٦ – ٢) و $s^2 = \sum e_i^2/(n-2) \cong 47.31/8 \cong 5.91$ ، (من مثال ٦ – ٣) .

فإذا توقعنا أن كمية السماد المستخدمة للفدان سنة ١٩٨١ سوف تكون $X_F = 35$ ، فإننا نحصل على

$$s_F^2 = 5.91\left[1 + \frac{1}{10} + \frac{(35-18)^2}{576}\right] \cong 9.46 \quad \text{and} \quad s_F \cong 3.08$$

$$\hat{Y}_F = 27.12 + 1.66(35) = 45.38$$

فتكون فترة الثقة 95 % أو فترة التنبؤ بالنسبة إلى Y_F فى عام ١٩٨١ هى $45.38 \pm (2.31)(3.08)$ ، أو بين 38.27 و 52.49 . (انظر المسألة ٨ – ٢٠ للتنبؤ فى حالة تحليل الانحدار المتعدد) .

مسائل محــلولة

شكل الــدالة :

٨ – ١ (أ) كيف يقرر شكل العلاقة الدالية ؟ (ب) ماهى بعض التحويلات إلى دوال خطية الأكثر فائدة ؟ (جـ) هل تكون المعالم المقدرة بتطبيق طريقة OLS على الدوال الخطية المحولة تقديرات غير متحيزة لمعالم المجتمع الحقيقية ؟

(أ) فى بعض الأحيان يمكن أن تقترح النظرية الاقتصادية شكل الدالة لعلاقة إقتصادية ما . فمثلا ، نظرية الاقتصاد الجزئى أن منحنى متوسط التكلفة (فى الأجل القصير) يأخذ شكل U وأن منحنى متوسط التكلفة الثابتة يتناقص باستمرار ويقترب – فى النهاية – من محور الكمية حيث يقسم إجمالى التكاليف الثابتة على عدد أكبر فأكبر من الوحدات المنتجة . كما قد يوحى شكل انتشار النقاط أيضاً بشكل الدالة المناسب فى حالة علاقة بين متغيرين . وعندما لايتوفر اقتراح بشكل العلاقة سواء عن طريق النظرية أو عن طريق شكل الانتشار ، فإنه عادة ما يتم تجربة الدالة الخطية لبساطتها .

(ب) بعض التحويلات الأكثر فائدة والأكثر شيوعاً من دوال غير خطية إلى دوال خطيه هي الدوال اللوغاريتمية المزدوجة ، نصف اللوغاريتمية ، المقلوبة ، كثيرات الحدود (انظر جدول ٨ – ١) . ومن مزايا الصورة اللوغاريتمية المزدوجة أن معالم الميل تمثل المرونات (أنظر المسألة ٨ – ٢) . وتكون الدالة نصف اللوغاريتمية ملائمة عندما يزيد المتغير التابع بمعدل ثابت تقريباً مع الزمن ، كما فى حالة القوة العاملة والسكان (أنظر المسألة ٨ – ٤) . أما الدالة المقلوبة والدالة كثيرة الحدود فتعتبر ملائمة لتقدير منحنيات متوسط التكاليف وإجمالى التكاليف (أنظر المسألة ٨ – ٥) .

(ج) يؤدى تقدير الدالة اللوغاريتمية المزدوجة المحولة باستخدام طريقة OLS إلى مقدرات للميل غير متحيزة . ولكن ، $\hat{b}_0 =$ العدد المقابل للوغاريتم للمعامل b_0 يكون مقدراً متحيزاً ، وإن كان متسقاً ، للمعلمة b_0 . وحقيقة أن $\hat{b}_0$ تكون متحيزة ليس لها تأثير كبير حيث أن الثابت عادة لايكون محل اهتمام أساسى (أنظر مسألة ٧ – ٦ (هـ)) . $\hat{b}_0$ غير متحيزة أيضاً فى الدوال المحلولة الأخرى فى جدول ٨ – ١) . ويكون النموذج اللوغاريتمى المزدوج الخطى مناسباً إذا وقعت النقاط لو Y – لو X تقريباً على خط مستقيم .

٨ – ٢ أثبت أنه فى معادلة الطلب اللوغاريتمية المزدوجة على الصورة

$$Q = b_0 P^{b_1} Y^{b_2} e^u$$

حيث $Q =$ الكمية المطلوبة ، $P =$ السعر ، $Y =$ الدخل ، (أ) b_1 هى المرونة السعرية للطلب ، أو η_P ، (ب) b_2 هى المرونة الدخلية للطلب أو η_Y . (يمكن للقارىء بدون العام بالتفاضل أن يتخطى هذه المسألة) .

(أ) تعريف المرونة السعرية للطلب هو :

$$\eta_P = \frac{dQ}{dP} \cdot \frac{P}{Q}$$

مشتقة دالة Q بالنسبة إلى P هى :

$$\frac{dQ}{dP} = b_1(b_0 P^{b_1 -1} Y^{b_2} e^u) = b_1(b_0 P^{b_1} Y^{b_2} e^u) P^{-1} = b_1 \cdot \frac{Q}{P}$$

بإحلال قيمة dQ/dP فى صيغة η_P نحصل على :

$$\eta_P = \frac{dQ}{dP} \cdot \frac{P}{Q} = b_1 \cdot \frac{Q}{P} \cdot \frac{P}{Q} = b_1$$

(ب) تعريف المرونة الدخلية للطلب هو :

$$\eta_Y = \frac{pQ}{dY} \cdot \frac{Y}{Q}$$

مشتقة الدالة Q بالنسبة إلى Y هى :

$$\frac{dQ}{dY} = b_2(b_0 P^{b_1} Y^{b_2 - 1} e^u) = b_2(b_0 P^{b_1} Y^{b_2} e^u) Y^{-1} = b_2 \cdot \frac{Q}{Y}$$

بإحلال قيمة dQ/dY فى صيغة η_Y ، نحصل على :

$$\eta_Y = \frac{dQ}{dY} \cdot \frac{Y}{Q} = b_2 \cdot \frac{Q}{Y} \cdot \frac{Y}{Q} = b_2$$

٨ – ٣ يعطى جدول (٨ – ٤) الإنتاج بالأطنان ، Q ، ومدخلات العمل بالساعة ، L ، ومدخلات رأس المال ، ساعة ماكينة ، K ، لعدد 14 شركة فى صناعة ما . والمطلوب توفيق البيانات لدالة إنتاج كوب – دوجلاس .

$$Q = b_0 L^{b_1} K^{b_2} e^u$$

جدول (٨ – ٤) الإنتاج ومدخلات العمل ورأس المال فى 14 شركة فى صناعة ما

الشركة	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Q	240	400	110	530	590	470	450	160	290	490	350	550	560	430
L	1,480	1,660	1,150	1,790	1,880	1,860	1,940	1,240	1,240	1,850	1,570	1,700	2,000	1,850
K	410	450	380	430	480	450	490	395	430	460	435	470	480	440

تحول البيانات أولا إلى الصورة اللوغاريتمية الطبيعية ، كما هو موضح فى جدول (٨ – ٥) ، ثم تطبق طريقة OLS على المتغيرات المحلولة كما سبق شرحه فى قسم (٦ – ٢) (ويقوم الكمبيوتر بكل هذا) . النتائج هى

$$\ln Q = -23.23 + \underset{(2.55)}{1.43} \ln L + \underset{(2.23)}{3.05} \ln K \qquad R^2 = 0.88$$

وتشير المعاملات المقدرة 1.43 و3.05 على الترتيب ، إلى مرونة الإنتاج بالنسبة إلى L وK . وحيث أن $1.43 + 3.05 = 4.48 > 1$ ، فإن هناك تزايداً فى الغلة لهذه الصناعة (بمعنى أن زيادة مدخلات كل من L وK بمقدار 10% يؤدى إلى زيادة الإنتاج بمقدار 44.8%) .

جدول (٨ – ٥) الإنتاج ومدخلات العمل ورأس المال فى صورتها الأصلية وفى الصورة اللوغاريتمية

الشركة	Q	L	K	$\ln Q$	$\ln L$	$\ln K$
1	240	1480	410	5.48064	7.29980	6.01616
2	400	1660	450	5.99146	7.41457	6.10925
3	110	1150	380	4.70048	7.04752	5.94017
4	530	1790	430	6.27288	7.48997	6.06379
5	590	1880	480	6.38012	7.53903	6.17379
6	470	1860	450	6.15273	7.52833	6.10925
7	450	1940	490	6.10925	7.57044	6.19441
8	160	1240	395	5.07517	7.12287	5.97889
9	290	1240	430	5.66988	7.12287	6.06379
10	490	1850	460	6.19441	7.52294	6.13123
11	350	1570	435	5.85793	7.35883	6.07535
12	550	1700	470	6.30992	7.43838	6.15273
13	560	2000	480	6.32794	7.60090	6.17379
14	430	1850	440	6.06379	7.52294	6.08677

٨ – ٤ يعطى جدول (٨ – ٦) عدد الأشخاص العاملين ، N ، (إلى أقرب مليون) فى الولايات المتحدة من عام ١٩٦٠ إلى ١٩٧٩ . المطلوب توفيق خط انحدار OLS إلى بيانات جدول (٨ – ٦) .

جدول (٨ – ٦) عدد الأشخاص العاملين فى الولايات المتحدة من ١٩٦٠ إلى ١٩٧٩

السنة	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969
N	69.6	70.5	70.6	71.8	73.1	74.5	75.8	77.3	78.7	80.7
السنة	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
N	82.7	84.1	86.5	88.7	91.0	92.6	94.8	97.4	100.4	102.9

المصدر : التقرير الاقتصادى للرئيس ، مكتب حكومة الولايات المتحدة للطباعة ، واشنطن ١٩٨٠ صفحة ٢٣٤ .

حيث أن التوظيف يميل إلى الزيادة بمعدل ثابت مع الزمن ، T ، فإنه يمكننا توفيق دالة نصف لوغاريتمية على الصورة الواردة فى معادلة (٨ – ٢) للبيانات المحولة فى جدول (٨ – ٧) . النتائج هى

$$\ln N = 4.19 + \underset{(41.18)}{0.02T} \qquad R^2 = 0.99$$

جدول (٨ – ٧) عدد العاملين فى الولايات المتحـــدة

١٩٦٠ – ١٩٧٩ : البيانات الأصلية والمحولة

السنة	N	$\ln N$	T
1960	69.6	4.24276	1
1961	70.5	4.25561	2
1962	70.6	4.25703	3
1963	71.8	4.27388	4
1964	73.1	4.29183	5
1965	74.5	4.31080	6
1966	75.8	4.32810	7
1967	77.3	4.34769	8
1968	78.7	4.36564	9
1969	80.7	4.39074	10
1970	82.7	4.41522	11
1971	84.1	4.43201	12
1972	86.5	4.46014	13
1973	88.7	4.48526	14
1974	91.0	4.51086	15
1975	92.6	4.52829	16
1976	94.8	4.55177	17
1977	97.4	4.57883	18
1978	100.4	4.60916	19
1979	102.9	4.63376	20

٨ – ٥ المطلوب توفيق منحنى متوسط التكلفة قصير الأجل لبيانات جدول (٨ – ٨) ، الذى يعطى متوسط التكلفة ، AC ، والإنتاج Q لشركة ما خلال فترة 12 أسبوعاً .

جدول (٨ – ٨) متوسط التكاليف والإنتـــاج لشركة ما خلال فترة 12 أسبوعاً

الأسبوع	1	2	3	4	5	6	7	8	9	10	11	12
AC	82	86	100	100	95	85	110	88	86	108	87	87
Q	149	121	190	100	109	138	209	170	158	201	130	181

حيث أن نظرية الاقتصاد الجزئى تفترض منحنى تكاليف للأجل القصير على شكل U ، فإننا نعمل على توفيق

$$AC = b_0 - b_1 Q + b_2 W + u \quad \text{حيث} \quad W = Q^2$$

وتكون النتيجة

$$\widehat{AC} = 244.86 - \underset{(-9.84)}{2.20Q} + \underset{(10.42)}{0.01Q^2} \qquad R^2 = 0.94$$

المتغيرات الصورية :

٨ – ٦ (أ) اكتب معادلة لوقت السلم وأخرى لوقت الحرب للمعادلات (٨ – ٥) ، (٨ – ٦) ، و (٨ – ٧) ، إذا كانت $C =$ الاستهلاك ، $Y_d =$ الدخل المتاح و $D = 1$ فى سنوات الحرب ، $D = 0$ فى سنوات السلم .

(ب) ارسم شكلا لمعادلات (٨ – ٥) ، (٨ – ٦) ، و (٨ – ٧) مبيناً دالة أستهلاك فى سنوات السلم ، ودالة استهلاك فى سنوات الحرب

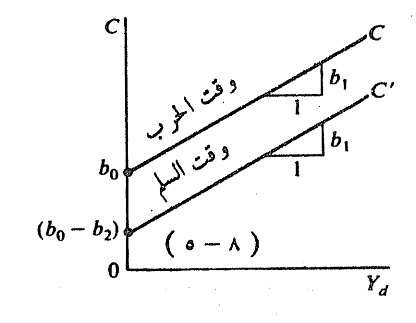
(جـ) ماهى مزايا تقدير المعادلات (٨ – ٥) ، (٨ – ٦) ، و (٨ – ٧) بدلا من تقدير انحدارين ، أحدهما لسنوات السلم والآخر لسنوات الحرب ، فى كل حالة ؟

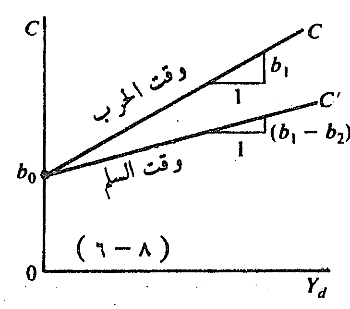
(أ) بجعل المعادلات التى تحتوى على a تشير إلى وقت السلم ، والمعادلات التى تحتوى على b تشير إلى وقت الحرب ، نحصل على

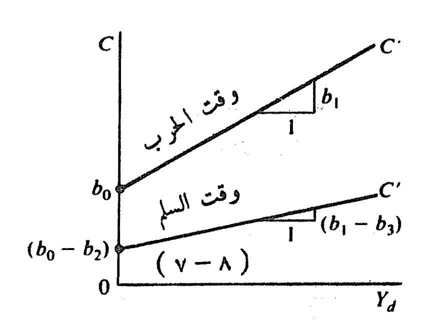
$$C = b_0 + b_1 Y_d + u \qquad (٨ – ٥ أ)$$
$$C' = (b_0 + b_2) + b_1 Y_d + u \qquad (٨ – ٥ ب)$$
$$C = b_0 + b_1 Y_d + u \qquad (٨ – ٦ أ)$$
$$C' = b_0 + (b_1 + b_2) Y_d + u \qquad (٨ – ٦ ب)$$
$$C = b_0 + b_1 Y_d + u \qquad (٨ – ٧ أ)$$
$$C' = (b_0 + b_2) + (b_1 + b_3) Y_d + u \qquad (٨ – ٧ ب)$$

لاحظ أن كل معادلات وقت السلم متطابقة لأن $D = 0$. خلال وقت الحرب ، يكون الاستهلاك أقل من وقت السلم بسبب القيود ونقص المتاح من السلع والخدمات ، وبسبب الدوافع الخلقية . وعليه فإن b_2 و b_3 (معاملات D) من المتوقع أن تكون سالبة لسنوات الحرب ، بحيث يكون المقطع و / أو الميل أقل لمعادلات سنوات الحرب عن معادلات سنوات السلم .

(ب) انظر شكل (٨ – ١)







(شكل ٨ – ١)

(جـ) إن مزايا تقدير معادلات (٨ – ٥) ، (٨ – ٦) ، و (٨ – ٧) بالمقارنة مع تقدير معادلة المقطع منفصلة لكل حالة ، واحدة لوقت السلم وأخرى لوقت الحرب هى (١) درجات حرية أكبر (٢) يمكن بسهولة اختبار عدد من الفروض لمعرفة ما إذا كانت الفروق فى الثوابت و / أو معاملات الميل معنوية إحصائياً ، (جـ) توفير وقت الكمبيوتر .

٨ – ٧ يعطى جدول (٨ – ٩) كمية الألبان (بالألاف (الواحدة ¼ جالون)) التى توردها شركة ما خلال شهر Q عند أسعار مختلفة ، P ، على فترة زمنية 14 شهراً . وقد واجهت الشركة إضراباً فى بعض مصانعها خلال الشهور الخامس ، والسادس ،

والسابع . بإجراء انحدار Q على P المطلوب (أ) إختبار ما إذا كان هناك تحرك فى المقطع خلال فترات الإضراب وعدم الإضراب (ب) إختبار ما إذا كان هناك تحرك فى المقطع وفى الميل .

جدول (٨ — ٩) الكمية الموردة من الألبان (بالآلاف) عند أسعار مختلفة

الشهر	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Q	98	100	103	105	80	87	94	113	116	118	121	123	126	128
P	0.79	0.80	0.82	0.82	0.93	0.95	0.96	0.88	0.88	0.90	0.93	0.94	0.96	0.97

(أ) بوضع $D = 1$ خلال شهور الإضراب و $D = 0$ فى غير ذلك ، نحصل على

$$\hat{Q} = -32.47 + \underset{(15.65)}{165.97P} - \underset{(-23.59)}{37.64D} \qquad R^2 = 0.98$$

حيث D معنوية إحصائياً عند مستوى أفضل من 1% ، فالمقطع هو $\hat{b}_0 = -32.47$ خلال فترات عدم الإضراب ، ويساوى $\hat{b}_0 + \hat{b}_2 = 32.47 - 37.64 = -70.11$ خلال فترة الإضراب .

(ب)

$$\hat{Q} = -29.74 + \underset{(27.16)}{162.86P} - \underset{(-5.67)}{309.62D} + \underset{(4.98)}{287.14XD} \qquad R^2 = 0.99$$

كل من D و XD معنوية إحصائياً عند مستوى أفضل من 1% . ويكون المقطع والميل على الترتيب ، هما 29.74 — و 162.86 خلال فترات عدم الإضراب . أما فى فترات الإضراب فإن المقطع يساوى

$$\hat{b}_0 + \hat{b}_2 = -29.74 - 309.26 = -339$$

بينما يكون الميل $\hat{b}_1 + \hat{b}_3 = 162.86 + 287.14 = 450$ (وحيث يفترض أن الشركة تستطيع زيادة إنتاجها بدرجة كبيرة فى المصانع التى ليس بها إضراب) .

٨ — ٨ يعطى جدول (٨ — ١٠) الإنفاق الاستهلاكى ، C ، والدخل المتاح ، Y_d ، وجنس رب البيت S لعدد 12 عائلة عشوائية . (أ) أوجد انحدار C على Y_d . (ب) اختبر الفرق فى المقطع للعائلات التى ربها رجل وتلك التى ربتها سيدة (جـ) اختبر الفرق فى العمل أو الميل الحدى للاستهلاك MPC للعائلات التى ربها رجل دن العائلات التى ربتها سيدة . (د) اختبر الفرق فى كل من المقطع والميل (هـ) أى النتائج هى « الأفضل » ؟

جدول (٨ — ١٠) الاستهلاك ، الدخل المتاح ، وجنس رب البيت لإثنتاعشرة أسرة عشوائية

العائلة	1	2	3	4	5	6	7	8	9	10	11	12
C	18,535	11,350	12,130	15,210	8,680	16,760	13,480	9,680	17,840	11,180	14,320	19,860
Y_d	22,550	14,035	13,040	17,500	9,430	20,635	16,470	10,720	22,350	12,200	16,810	23,000
S	M	M	F	M	F	M	M	F	M	F	F	M

(أ)

$$\hat{C} = \underset{(2.73)}{1,663.60} + \underset{(21.12)}{0.75Y_d} \qquad R^2 = 0.978$$

(ب) بوضع $D = 1$ للعائلات التى على رأسها سيدة و $D = 0$ لغير ذلك ، نحصل على

$$\hat{C} = 186.12 + \underset{(16.56)}{0.82\,Y_d} + \underset{(1.82)}{832.09\,D} \qquad R^2 = 0.984$$

(جـ)

$$\hat{C} = 709.18 + \underset{(18.11)}{0.79\,Y_d} + \underset{(1.51)}{0.05\,Y_d D} \qquad R^2 = 0.983$$

(د)

$$\hat{C} = -184.70 + \underset{(13.65)}{0.83\,Y_d} + \underset{(1.03)}{1{,}757.99\,D} - \underset{(-0.57)}{0.06\,Y_d D} \qquad R^2 = 0.985$$

(و) حيث أن كلا من D ، $D\,Y_d$ غير معنوية إحصائياً عند مستوى 5% فى أجزاء (ب) ، (جـ) ، و (د) ، فإنه لايوجد اختلاف فى نمط الاستهلاك بين العائلات التى يرأسها رجال وتلك التى ترأسها سيدات . وتكون أفضل النتائج تلك الواردة فى (أ) .

٨ — ٩ يعطى جدول (٨ — ١١) الأرباح (بعد الضرائب) والمبيعات للمؤسسات الإنتاجية الأمريكية من الربع الأول لعام ١٩٧٤ حتى الربع الثالث لعام ١٩٧٩ . (أ) رتب جدولا يوضح الأرباح والمبيعات ومتغيراً صورياً ليأخذ فى الحساب الآثار الموسمية (ب) باستخدام البيانات من الجدول فى (أ) أجر تقدير الانحدار للأرباح على المبيعات والمتغيرات الصورية الموسمية وفسر النتائج .

جدول (٨ — ١١) الأرباح والمبيعات للمؤسسات الإنتاجية فى الولايات المتحدة (ببلايين الدولارات)

الأرباح	13.5	16.3	15.5	13.4	9.3	12.4	13.2	14.2	14.8	18.1	16.0	15.6
المبيعات	242.0	269.4	272.1	277.0	247.1	265.8	271.0	281.3	284.2	307.6	301.6	309.8
الربع	I	II	III	IV	I	II	III	IV	I	II	III	IV
السنة	1974				1975				1976			

الأربـــاح	15.6	19.7	16.7	18.4	16.0	22.1	20.4	22.6	22.6	26.8	24.8
المبيعات	311.5	338.6	331.7	346.2	340.2	377.5	376.9	401.8	406.2	436.4	437.5
الربـــع	I	II	III	IV	I	II	III	IV	I	II	III
السنة	1977				1978				1979		

المصدر : التقرير الاقتصادى للرئيس ، مكتب الحكومة الأمريكية للطباعة ، واشنطن ، ١٩٨٠ صفحة ٣٠٠

(أ) بأخذ الربع الأول كأساس ، ووضع $D_1 = 1$ للربع الثانى وتساوى الصفر لغير ذلك . $D_2 = 1$ للربع الثالث وتساوى الصفر لغير ذلك ، و $D_3 = 1$ للربع الرابع وتساوى الصفر لغير ذلك ، فإننا نحصل على جدول (٨ — ١٢) .

(ب) باستخدام بيانات جدول (٨ — ١٢) لإيجاد انحدار الأرباح ، P ، على المبيعات S ، D_1 ، D_2 و D_3 نحصل على

$$\hat{P} = -5.20 + \underset{(14.79)}{0.07\,S} + \underset{(2.91)}{2.10\,D_1} + \underset{(0.94)}{0.68\,D_2} + \underset{(0.44)}{0.33\,D_3} \qquad R^2 = 0.93$$

وحيث أن D_1 فقط ذات معنوية إحصائية عند مستوى 5% ،

فى الربع الأول والثالث والرابع $\hat{P} = -5.20 + 0.07\,S$

فى الربـــع الثـــانى $\hat{P} = -3.10 + 0.07\,S$

ولا تتغير هذه النتائج لو استخدمنا أربعة متغيرات صورية ، واحد لكل ربع سنة ، **ولكن** يتم حذف ثابت معادلة الانحدار . أن استخدام المتغيرات الصورية الأربعة مع الثابت يجعل من غير الممكن تقدير انحدار OLS (أنظر قسم ٩ — ٢) .

جدول (٨ – ١٢) الأرباح ، المبيعات ، والمتغيرات الموسمية الصورية

المبيعات	الأرباح	الربع	السنة	D_1	D_2	D_3
1974	I	13.5	242.0	0	0	0
	II	16.3	269.4	1	0	0
	III	15.5	272.1	0	1	0
	IV	13.4	277.0	0	0	1
1975	I	9.3	247.1	0	0	0
	II	12.4	265.8	1	0	0
	III	13.2	271.0	0	1	0
	IV	14.2	281.3	0	0	1
1976	I	14.8	284.2	0	0	0
	II	18.1	307.6	1	0	0
	III	16.0	301.6	0	1	0
	IV	15.6	309.8	0	0	1
1977	I	15.6	311.5	0	0	0
	II	19.7	338.6	1	0	0
	III	16.7	331.7	0	1	0
	IV	18.4	346.2	0	0	1
1978	I	16.0	340.2	0	0	0
	II	22.1	377.5	1	0	0
	III	20.4	376.9	0	1	0
	IV	22.6	401.8	0	0	1
1979	I	22.6	406.2	0	0	0
	II	26.8	436.4	1	0	0
	III	24.8	437.5	0	1	0

٨ – ١٠ (أ) ماذا يقصد بمتغير تابع كيفى ؟ (ب) ماهى المشاكل أو الصعوبات التى تنشأ فى انحدار ذى متغير تابع كيفى ؟

(أ) المتغير التابع الكيفى : هو متغير ثنائى الوجه ، مشيراً إلى حدوث أو عدم حدوث حدث ما أو إلى وجود أو غياب ظروف معينة . فمثلا ، يكون الشخص داخل القوة العاملة أو خارجها ، عاملا أو غير عامل . يكون لدى الشخص سيارة ومنزل ويذهب للجامعة أولا . فى هذه الحالات ، فإن وقوع الحدث ، أو وجود الظروف تبين له عادة القيمة 1 ، بينما يعطى عدم الحدوث أو الغياب القيمة 0 .

(ب) عندما يكون المتغير التابع ثنائى الوجه أو كيفى ، يظل من الممكن استخدام طريقة المربعات الصغرى لتقدير معادلة الانحدار ولكن تنشأ بعض المشاكل . الأولى ، مخالفة فرض نموذج الانحدار الخطى الكلاسيكى OLS من أن حد الخطأ u يتبع التوزيع الطبيعى (أنظر قسم ٦ – ١) . وهذا الفرض مطلوب لإجراء اختبارات معنوية المعالم . على أنه يمكن اللجوء إلى نظرية النهاية المركزية (أنظر قسم ٤ – ٢) للعينات الكبيرة $(n \geq 30)$. الثانية ، مخالفة الفرض بأن عنصر الخطأ غير مرتبط بالمتغير المفسر . ؟ ويؤدى هذا إلى مقدرات غير متحيزة ولكن غير كفوء (أى أن المقدرات لايكون لها أصغر تباين) . ولكن يمكن أيضاً التغلب على هذه المشكلة كما هو موضح فى قسم(٩ – ٢) . وأخيراً ، فإن القيم المقدرة للمتغير التابع قد تأخذ قيما خارج المدى من 0 إلى 1 . ويمكن التغلب على هذه الصعوبة بضغط الاحتمالات المقدرة داخل المدى من 0 إلى 1 أما باستخدام الدالة الطبيعية التراكية (نموذج probit) أو الدالة اللوجتية (نموذج logit) . وتقدم أساليب التقدير هذه فى كتب الاقتصاد القياسى الأكثر تقدماً .

نماذج فترات الإبطاء الموزعة :

٨ – ١١ (أ) ماذا يقصد بنموذج فترات الإبطاء الموزعة ؟ (ب) اكتب المعادلة للنموذج العام لفترات الإبطاء الموزعة مع عدد لانهائى من فترات الإبطاء وأخرى لعدد k من فترات الإبطاء . (جـ) ماهى الصعوبات العملية التى تنشأ من تقدير نموذج بعدد k من فترات الإبطاء ؟

(أ) غالباً مايتوزع تأثير متغير يتعلق بالسياسات على سلسلة من الفترات الزمنية (أى أن المتغير التابع قد يكون بطيء الاستجابة للتغير فى السياسات) مما يتطلب سلسلة من التغيرات المفسرة المبطأة لتفسير عملية التكيف الكاملة خلال الزمن . **نموذج فترات الإبطاء الموزعة** هو نموذج تعتمد فيه القيمة الحاضرة للمتغير التابع ، Y_t ، على المجموع المرجح لقيم الحاضرة والماضية للمتغيرات المستقلة (أى X_t ، X_{t-1} ، X_{t-2} ، الخ) وعلى حد الخطأ ، مع تعيين أوزان مختلفة عدة للفترات الزمنية المختلفة (عادة تتناقص مع تباعد الفترات الزمنية) .

(ب) (٨ – ٩) $$Y_t = a + b_0X_t + b_1X_{t-1} + b_2X_{t-2} + \cdots + u_t$$

(٨ – ٩ أ) $$Y_t = a + b_0X_t + b_1X_{t-1} + b_2X_{t-2} + \cdots + b_kX_{t-k} + u_t$$

لاحظ أنه فى معادلات (٨ – ٩) و (٨ – ٩ أ) ، a ثابت ، بينما b_0 هى معامل X_t . وقد فعلنا ذلك لتبسيط المعالجة الجبرية فى المسألة (٨ – ١٢ (أ) .

(جـ) عند تقدير نموذج فترات الإبطاء الموزعة ، فإنه فى مقابل إدخال حد مبطأ تفقد واحدة من درجات الحرية . وعندما يكون عدد الحدود المستقلة المبطأة ، k ، صغيراً فإنه يمكن تقدير النموذج باستخدام OLS ، كما فى الفصل السابع . ولكن ، عندما تكون k كبيرة (بالنسبة لطول السلسلة الزمنية) ، فإنه قد يتبقى عدد غير كاف من درجات الحرية لتقدير النموذج أو لكى يكون الإنسان واثقاً فى المعالم المقدرة . ثانياً ، فإن المتغيرات المفسرة المبطأة فى نموذج فترات الإبطاء الموزعة ، سوف يكون بينها على الأرجح ارتباط قوى ، وبالتالى قد يكون من الصعب الفصل بدرجة كافية بين تأثيرات المتغيرات المستقلة على المتغير التابع (انظر المسألة ٧ – ٣) (ب)) .

٨ – ١٢ (أ) اشتق نموذج فترات الإبطاء الموزعة لكويك. (ب) ما هى المشاكل التى تنشأ عند تقدير هذا النموذج ؟ (إرشاد بالنسبة لجزء (أ) : ابدأ بالنموذج العام لفترات الإبطاء الموزعة وافترض أن الأوزان تتناقص بمتوالية هندسية ، حيث تشير λ إلى ثابت أكبر من 0 وأصغر من 1 ، ثم ابطيء العلاقة بفترة زمنية واحدة ، واضرب فى λ ، واطرح من العلاقة الأصلية) .

(أ) بدءاً بمعادلة (٨ – ٩) ، من المفترض أن كل شروط OLS متوفرة (أنظر المسألة ٧ – ١) :

(٨ – ٩) $$Y_t = a + b_0X_t + b_1X_{t-1} + b_2X_{t-2} + \cdots + u_t$$

استخدام أوزان تتناقص كمتوالية هندسية و $0 < \lambda < 1$ يعطى

(٨ – ١٦) $$b_i = \lambda^i b_0 \qquad i = 1, 2, \ldots$$

بالتعويض من معادلة (٨ – ١٦) فى معادلة (٨ – ٩) .

$$Y_t = a + b_0X_t + \lambda b_0X_{t-1} + \lambda^2 b_0X_{t-2} + \cdots + u_t$$

بالإبطاء بفترة زمنية واحدة

$$Y_{t-1} = a + b_oX_{t-1} + \lambda b_0X_{t-2} + \lambda^2 b_0X_{t-3} + \cdots + u_{t-1}$$

بالضرب فى

$$\lambda Y_{t-1} = \lambda a + \lambda b_0X_{t-1} + \lambda^2 b_0X_{t-2} + \cdots + \lambda u_{t-1}$$

بالطرح من معادلة (٨ – ٩) ،

$$Y_t - \lambda Y_{t-1} = a - \lambda a + b_0X_t + \lambda b_0X_{t-1} - \lambda b_0X_{t-1} + \lambda^2 b_0X_{t-2} - \lambda^2 b_0X_{t-2} + \cdots + u_t - \lambda u_{t-1}$$

$$Y_t - \lambda X_{t-1} = a(1-\lambda) + b_0X_t + u_t - \lambda u_{t-1}$$

(٨ – ١٠) $$Y_t = a(1-\lambda) + b_0X_t + \lambda Y_{t-1} + v_t$$

حيث $\upsilon_t = u_t - \lambda u_{t-1}$

لاحظ أنه فى معادلة (٨ – ١٠) نقص عدد المتغيرات المنحدرة إلى اثنين ، مع بقاء X واحدة فقط .

(ب) تنشأ مشكلتان خطيرتان فى تقدير نموذج فترات الإبطاء الموزعة لكويك الأولى ، إذا كانت u_t فى معادلة (٨ – ٩) تتوفر فيها كل فروض OLS (أنظر المسألة ٦ – ٤) ، فإن $\upsilon_t = u_t - \lambda u_{t-1}$ فى معادلة (٨ – ١٠) لاتتوفر فيها هذه الشروط . وبالتحديد فإن ، $E(\upsilon_t\, \upsilon_{t-1}) \neq 0$ لأن كلا من υ_t و υ_{t-1} كما تتضمن فى تعريفها القيمة u_{t-1} (أى $\upsilon_t = u_t - \lambda u_{t-1}$ و $\upsilon_{t-1} = u_{t-1} - \lambda u_{t-2}$ بالإضافة إلى أن ، $E(\upsilon_t\, Y_{t-1}) \neq 0$. أن خرق فروض OLS هذه ينتج عنها مقدرات متحيزة وغير متسقة لنموذج كويك المبطأ (معادلة (٨ – ١٠)) ، وتتطلب إجراءات تصحيحية مفصلة (يناقش بعضها فى قسم ٩ – ٣) . المشكلة الخطيرة الثانية هى أن نموذج كويك يفترض أوزاناً متناقصة كمتوالية هندسية . وغالباً مايكون الواقع غير ذلك ، متطلباً بذلك خطة إبطاء أكثر مرونة (أنظر المسألة ٨ – ١٤) .

٨ – ١٣ يعطى جدول (٨ – ١٣) مستوى المخزون ، Y ، والمبيعات ، X ، ببلايين الدولارات ، فى الصناعة التمويلية الأمريكية من ١٩٥٩ إلى ١٩٧٨ والمطلوب (أ) توفيق نموذج كويك لبيانات جدول (٨ – ١٣) (ب) ماقيمة λ و a ؟

جدول (٨ – ١٣) المخزون والمبيعات فى الصناعات التمويلية الأمريكية ، ١٩٥٩–١٩٧٨
(ببلايين الدولارات)

السنة	1959	1960	1961	1962	1963	1964	1965	1966	1967	1968
Y	52.9	53.8	54.9	58.2	60.0	63.4	68.2	78.0	84.7	90.6
X	30.3	30.9	30.9	33.4	35.1	37.3	41.0	44.9	46.5	50.3
السنة	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978
Y	98.2	101.7	102.7	108.3	124.7	157.9	158.2	170.2	180.0	198.0
X	53.5	52.8	55.9	63.0	73.0	84.8	86.6	98.9	110.8	124.7

المصدر : التقرير الاقتصادى للرئيس ، مكتب حكومة الولايات المتحدة للطباعة واشنطن ١٩٨٠ ، صفحة ٢٥٦

(أ)
$$\hat{Y}_t = 2.37 + 0.94 X_t + 0.47 Y_{t-1} \qquad R^2 = 0.99$$
$$\qquad\qquad (4.17) \qquad (3.15)$$

(ب)
$$\hat{\lambda} = 0.47 \quad \text{and} \quad \hat{a}(1 - 0.47) = 2.37 \quad \text{so} \quad \hat{a} = 4.47$$

٨ – ١٤ (أ) ماهو هيكل الإبطاء فى نموذج ألمون للإبطاء ؟ (ب) ماهى مزايا وعيوب نموذج ألمون للإبطاء بالمقارنة مع نموذج كويك ؟

(أ) بينما يفترض نموذج كويك للابطاء أوزاناً متناقصة بمتوالية هندسية ، فإن نموذج ألمون للابطاء يسمح بأى هيكل للابطاء ، على أن يقرب عملياً بكثيرة حدود درجتها تزيد عن عدد نقاط التحول فى الدالة واحداً . فمثلا ، هيكل إبطاء على صورة U معكوسة (أى عندما $b_1 > b_0$) يمكن تقريبها باستخدام كثيرة حدود من الدرجة الثانية على الأقل . وقد ينشأ هذا ، كما فى حالة دالة الاستثمار ، وبسبب التأخر الناشىء عن تفهم الظروف واتخاذ القرارات ، فإن مستوى الاستثمار فى الفترة الحالية يكون أكثر استجابة لظروف الطلب فى الفترات القريبة الماضية عنه فى الفترة الحالية .

(ب) لنموذج ألمون للابطاء ميزتان هامتان عن نموذج كويك للابطاء . الأولى (كما أشرنا سابقاً) ، أن لنموذج ألمون هيكل إبطاء مرن على عكس هيكل الإبطاء الجامد لنموذج كويك . الثانية ، أنه حيث أن نموذج ألمون للابطاء لايستبدل متغيراً تابعاً مبطئاً بالمتغيرات المستقلة المبطأة ، فإنه لايخرق أياً من فروض OLS (كما يفعل نموذج كويك) . ومن عيوب

نموذج ألمون أن عدد المعاملات اللازم تقديرها لاينخفض كثيراً كما يحدث فى نموذج كويك . ومن عيوبه أيضاً أنه فى الواقع العملى ، قد لانستطيع تحديد فترة الإبطاء أو شكل الإبطاء عن طريق النظرية أو بمعلومات مسبقة

٨ – ١٥ اشتق تحويلة ألمون لكل من (أ) ثلاث فترات إبطاء تأخذ شكل كثيرة حدود من الدرجة الثانية (ب) أربع فترات إبطاء تأخذ شكل كثيرة حدود من الدرجة الثالثة .

(أ) بدءاً بمعادلة (٨ – ١١) و (٨ – ١٢) :

$$Y_t = a + b_0X_t + b_1X_{t-1} + b_2X_{t-2} + b_3X_{t-3} + u_t \qquad (٨ - ١١)$$

$$b_i = c_0 + c_1 i + c_2 i^2 \quad \text{with } i = 0, 1, 2, 3 \qquad (٨ - ١٢)$$

وبالتعويض بمعادلة (٨ – ١٢) فى معادلة (٨ – ١١) ، نحصل على :

$$Y_t = a + c_0X_t + (c_0 + c_1 + c_2)X_{t-1} + (c_0 + 2c_1 + 4c_2)X_{t-2} + (c_0 + 3c_1 + 9c_2)X_{t-3} + u_t$$

وبإعادة ترتيب الحدود فى التعبير الأخير :

$$Y_t = a + c_0\left(\sum_{i=0}^{3} X_{t-i}\right) + c_1\left(\sum_{i=1}^{3} iX_{t-i}\right) + c_2\left(\sum_{i=1}^{3} i^2X_{t-i}\right) + u_t$$

بوضع $Z_{1t} = \sum_{i=0}^{3} X_{t-i}$, $Z_{2t} = \sum_{i=1}^{3} iX_{t-i}$, and $Z_{3t} = \sum_{i=1}^{3} i^2X_{t-i}$ نحصل على

$$Y_t = a + c_0Z_{1t} + c_1Z_{2t} + c_2Z_{3t} + u_t \qquad (٨ - ١٣)$$

(ب) باستخدام إبطاء لمدة أربع فترات فى شكل كثيرة حدود من الدرجة الثالثة يكون لدينا

$$Y_t = a + b_0X_t + b_1X_{t-1} + b_2X_{t-2} + b_3X_{t-3} + b_4X_{t-4} + u_t$$

$$b_i = c_0 + c_1 i + c_2 i^2 + c_3 i^3 \quad \text{with } i = 0, 1, 2, 3, 4$$

وبالتعويض من المعادلة الثانية فى المعادلة الأولى ، نحصل على .

$$Y_t = a + c_0X_t + (c_0 + c_1 + c_2 + c_3)X_{t-1} + (c_0 + 2c_1 + 4c_2 + 8c_3)X_{t-2}$$
$$+ (c_0 + 3c_1 + 9c_2 + 27c_3)X_{t-3} + (c_0 + 4c_1 + 16c_2 + 64c_3)X_{t-4} + u_t$$

وبإعادة ترتيب الحدود فى التعبير الأخير ، يكون لدينا

$$Y_t = a + c_0\left(\sum_{i=0}^{4} X_{t-i}\right) + c_1\left(\sum_{i=1}^{4} iX_{t-i}\right) + c_2\left(\sum_{i=1}^{4} i^2X_{t-i}\right) + c_3\left(\sum_{i=1}^{4} i^3X_{t-i}\right) + u_t$$

وبمساواة الحدود داخل الأقواس بالمقادير Z_{1t} ، Z_{2t} ، Z_{3t} و Z_{4t} على الترتيب نحصل على :

$$Y_t = a + c_0Z_{1t} + c_1Z_{2t} + c_2Z_{3t} + c_3Z_{4t} + u_t$$

٨ – ١٦ باستخدام بيانات جدول (٨ – ١٣) وبافتراض إبطاء لمدة ثلاث فترات فى صورة كثيرة حدود الدرجة الثانية :

(أ) رتب جدولا بالمتغيرات الأصلية وقيم Z المحسوبة لاستخدامها فى تقدير نموذج ألمون للابطاء .

(ب) أوجد انحدار مستوى المخزون ، Y ، على قيم Z الواردة فى الجدول (أ) ، أى قدر معادلة الانحدار (٨ – ١٣)

(جـ) أوجد قيم $\hat{b}$ واكتب المعادلة المقدرة (٨ – ١١) .

(أ) تحسب قيم Z المعطاة فى جدول (٨ - ١٤) كما يلى :

$$Z_{1t} = \sum_{i=0}^{3} X_{t-i} = (X_t + X_{t-1} + X_{t-2} + X_{t-3})$$

$$Z_{2t} = \sum_{i=1}^{3} iX_{t-i} = (X_{t-1} + 2X_{t-2} + 3X_{t-3})$$

$$Z_{3t} = \sum_{i=1}^{3} i^2X_{t-i} = (X_{t-1} + 4X_{t-2} + 9X_{t-3})$$

جدول (٨ - ١٤) المخزون ، المبيعات ، وقيم Z فى الصناعة التحويلية الأمريكية ، ١٩٦٠ - ١٩٧٩ (ببلايين الدولارات)

السنة	Y	X	Z_1	Z_2	Z_3
1960	52.9	30.3	—	—	—
1961	53.8	30.9	—	—	—
1962	54.9	30.9	—	—	—
1963	58.2	33.4	125.5	183.6	427.2
1964	60.0	35.1	130.3	187.9	435.1
1965	63.4	37.3	136.7	194.6	446.8
1966	68.2	41.0	146.8	207.7	478.3
1967	78.0	44.9	158.3	220.9	506.1
1968	84.7	46.5	169.7	238.8	544.6
1969	90.6	50.3	182.7	259.3	595.1
1970	98.2	53.5	195.2	278.0	640.4
1971	101.7	52.8	203.1	293.6	673.2
1972	102.7	55.9	212.5	310.7	719.6
1973	108.3	63.0	225.2	322.0	748.6
1974	124.7	73.0	244.7	333.2	761.8
1975	157.9	84.8	276.7	366.7	828.1
1976	158.2	86.6	307.4	419.8	943.8
1977	170.2	98.8	343.2	475.2	1,082.8
1978	180.0	110.8	381.0	526.4	1,208.4
1979	198.0	124.7	420.9	568.2	1,285.4

(ب) بإجراء انحدار Y على قيم Z ، نحصل على :

$$\hat{Y}_t = 8.68 + \underset{(1.94)}{0.91Z_{1t}} + \underset{(0.27)}{0.30Z_{2t}} - \underset{(-0.74)}{0.28Z_{3t}} \qquad R^2 = 0.98$$

(جـ)

$$\hat{a} = 8.68$$

$$\hat{b}_0 = \hat{c}_0 = 0.91$$

$$\hat{b}_1 = (\hat{c}_0 + \hat{c}_1 + \hat{c}_2) = (0.91 + 0.30 - 0.28) = 0.93$$

$$\hat{b}_2 = (\hat{c}_0 + 2\hat{c}_1 + 4\hat{c}_2) = (0.91 + 0.60 - 1.12) = 0.30$$

$$\hat{b}_3 = (\hat{c}_0 + 3\hat{c}_1 + 9\hat{c}_2) = (0.91 + 0.90 - 2.52) = -0.71$$

بحيث أن :

$$\hat{Y}_t = 8.68 + \underset{(1.94)}{0.91X_t} + \underset{(2.49)}{0.93X_{t-1}} + \underset{(0.28)}{0.30X_{t-2}} - \underset{(-1.05)}{0.71X_{t-3}}$$

حيث تم إيجاد الخطأ المعيارى لقيم X المبطأة كالآتى :

$$\sqrt{\text{var}\,\hat{b}_i} = \sqrt{\text{var}(\hat{c}_0 + \hat{c}_1 i + \hat{c}_2 i^2)} \qquad (٨ - ١٧)$$

التنبــؤ :

٨ – ١٧ (أ) ماذا يقصد بالتنبؤ ؟ **التنبؤ المشروط** ؟ **الإسقاط** ؟ (ب) ماهى مصادر الخطأ المحتملة فى التنبؤ ؟ (جـ) ماهوتباين خطأ التنبؤ؟ اذكر تقديراً غير متحيز لتباين خطأ التنبؤ . علام يعتمد ؟

(د) كيف يتم إيجاد قيمة $\hat{Y}_F$ ؟ فترة الثقة 95% للتنبؤ ، Y_F ؟

(أ) يشير التنبؤ إلى تقدير قيمة المتغير التابع ، بمعلومية القيمة الفعلية أو المتوقعة للمتغير المستقل (المتغيرات المستقلة) وعندما يبنى التنبؤ على قيمة مقدرة أو متوقعة (بدلا من قيمة فعلية) للمتغير المستقل ، يكون لدينا **تنبؤ مشروط** أما **الإسقاط** فإنه يستخدم بالتبادل مع التنبؤ . فى بعض الأحيان ، يستخدم **الإسقاط** عند تقدير قيمة المتغير التابع داخل العينة (فترة العينة) . ويشير **التنبؤ** فى هذه الحالة إلى تقدير قيمة مستقبلية للمتغير التابع .

(ب) ينشأ خطأ التنبؤ كنتيجة (١) للطبيعة العشوائية لحد الخطأ ، (٢) أن المعالم المقدرة غير المتحيزة تساوى المعالم الحقيقية فى المتوسط فقط (٣) أخطاء فى توقع المتغيرات المستقلة (٤) تحديد خاطىء للنموذج .

(جـ) تباين خطأ التنبؤ ، σ_F^2 هو :

$$\sigma_F^2 = \sigma_u^2\left[1 + \frac{1}{n} + \frac{(X_F - \bar{X})^2}{\sum(X_i - \bar{X})^2}\right] \qquad (٨ - ١٤)$$

حيث n عدد المشاهدات و σ_u^2 تباين u . وتقدير غير متحيز لتباين وخطأ التنبؤ ، s_F^2 هو :

$$s_F^2 = s^2\left[1 + \frac{1}{n} + \frac{(X_F - \bar{X})^2}{\sum(X_i - \bar{X})^2}\right] \qquad (٨ - ١٥)$$

حيث s^2 تقدير غير متحيز للتباين σ_u^2 ويحسب كالآتى :

$$s^2 = \frac{\sum(Y_i - \hat{Y}_i)^2}{n-2} = \frac{\sum e_i^2}{n-2} \qquad (٦ - ١٢)$$

كلما كبرت n ، صفر σ_F^2 (أو s_F^2) ، σ_u^2 (أو s^2) ، والفرق بين X_F و $\bar{X}$.

(د) يتم إيجاد قيمة $\hat{Y}_F$ بإحلال القيمة الفعلية أو المقدرة للمتغير X_F فى معادلة الانحدار المقدرة . أى

$$\hat{Y}_F = \hat{b}_0 + \hat{b}_1 X_F$$

فترة الثقة 95% للتنبؤ ، Y_F ، هى

$$\hat{Y}_F \pm t_{0.025} s_F$$

حيث تشير t إلى توزيع t بدرجات حرية $n - 2$.

٨ – ١٨ أوجد فترة الثقة 95% للتنبؤ بقيمة Y فى المسألة (٦ – ٣١) لكل من (أ) $X = 11\%$ (ب) X 8.5%

(أ) وجدنا فى المسألة (٦ – ٣٠) أن $\hat{Y}_i = 12.29 - 0.47 X_i$ ، $n = 15$ ، $\bar{X} = 7$ ، $\Sigma(X - \bar{X})^2 = 60$
$s^2 = 26.93/13 \cong 2.07$ عندما $X = 11\%$

$$s_F^2 = 2.07\left[1 + \frac{1}{15} + \frac{(11-7)^2}{60}\right] \cong 2.77 \qquad \text{and} \qquad s_F \cong 1.66$$

$$\hat{Y}_F = 12.29 - 0.47(11) = 7.12$$

وتكون فترة الثقة 95% للتنبؤ Y_F

$(2.18)(1.66) \pm 7.12$ أى بين 3.50 و 10.74 حيث $t_{0.025} = 2.18$ بدرجات حرية 13

(ب) عندما $X = 8.5\%$ ،

$$s_F^2 = 2.07\left[1 + \frac{1}{15} + \frac{(8.5-7)^2}{60}\right] \cong 2.30 \quad \text{and} \quad s_F \cong 1.52$$

$$\hat{Y}_F = 12.29 - 0.47(8.5) = 8.29$$

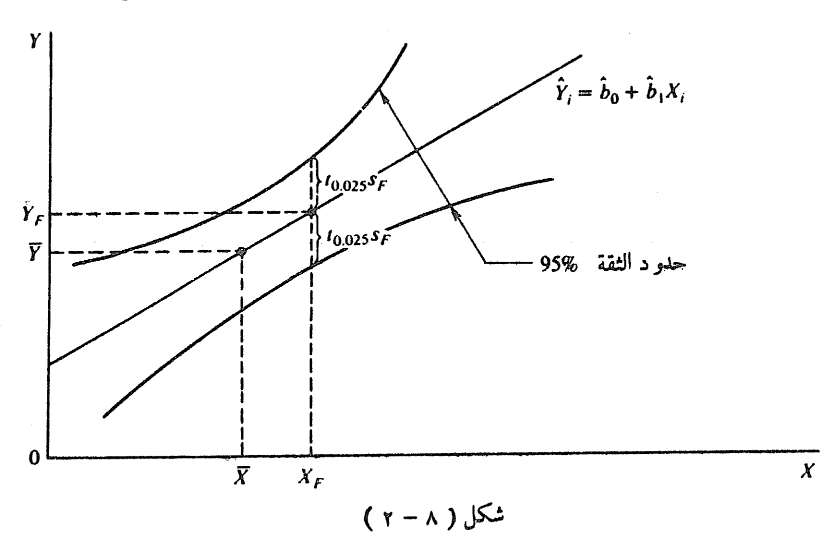
فتكون فترة الثقة 95% للتنبؤ Y_F

$(2.18)(1.62) \pm 8.29$ أى بين 4.98 و 11.60

لاحظ أن مدى فترة الثقة 95% للتنبؤ Y_F أصغر هنا عن مدى فترة الثقة فى (أ) لأن الفرق بين قيمة X المتوقعة وقيمة $\bar{X}$ أصغر هنا .

٨ – ١٩ ارسم شكلا يوضح خط تقدير انحدار OLS والمفترض أنه موجب الميل، وكذلك فترة الثقة 95% للتنبؤ Y_F عند قيمة معينة X_F وحدود فترة الثقة 95% للتنبؤ Y_F .

أنظر شكل (٨ – ٢) . لاحظ أن حدود فترة الثقة 95% أضيق ما تكون عند $X_F = \bar{X}$.



شكل (٨ – ٢)

٨ – ٢٠ أوجد فترة الثقة 95% للتنبؤ Y_F عند $X_{1F} = 35$ و $X_{2F} = 25$ فى ١٩٨١ ،

بمعلومية أن $\hat{Y}_i = 31.98 + 0.65 X_{1i} + 1.10 X_{2i}$ ، $\bar{X}_1 = 18$ ، $\bar{X}_2 = 12$

(من مثال ٧ – ١) ، $s^2 = \sum e_i^2/(n-k) = 13.67/7 \cong 1.95,\ s_{\hat{b}_1}^2 \cong 0.06,\ s_{\hat{b}_2}^2 \cong 0.07$

(من مثال ٧ – ٢) ، $7.2),\ s_{\hat{b}_0}^2 \cong 2.66,\ \operatorname{cov}(\hat{b}_1\hat{b}_2) \cong s_{\hat{b}_1\hat{b}_2} \cong 0.07$

(من حسابات الكمبيوتر) ، وإذا كانت

$$s_F^2 = s^2 + s_{\hat{b}_0}^2 + s_{\hat{b}_1}^2(X_{1F} - \bar{X}_1)^2 + s_{\hat{b}_2}^2(X_{2F} - \bar{X}_2)^2 + s_{\hat{b}_1\hat{b}_2}(X_{1F} - \bar{X}_1)(X_{2F} - \bar{X}_2) \quad (\text{٨ – ١٨})$$

$$s_F^2 = 1.95 + 2.66 + 0.06(35-18)^2 + 0.07(25-12)^2 + (-0.07)(35-18)(25-12)$$

$$\cong 18.31 \quad \text{and} \quad s_F \cong 4.28$$

$$\hat{Y}_F = 31.98 + 0.65(35) + 1.10(25) = 82.23$$

وتكون فترة الثقة 95% للتنبؤ Y_F فى عام ١٩٨١

$$82.23 \pm (2.37)(4.28)$$ أى بين 72.09 92.37

مسائل إضافية

شكل الدالة :

٨ — ٢١ حول الدوال غير الخطية التالية إلى دوال خطية : (أ) $Y = b_0 e^{b_1 X} e^u$ (ب) $Y = b_0 + b_1 \ln X + u$ (جـ) $Y = b_0 - b/X + u$ (د) $Y = b_0 + b_1 X - b_2 X^2 + b_3 X^3 + u$

الإجابة : (أ) $\ln Y = \ln b_0 + b_1 X + u$ (ب) $Y = b_0 + b_1 R + u$ حيث $R = \ln X$ ، (جـ) $Y = b_0 - b_0 Z + u$ حيث $Z = 1/X$ ، (د) $Y = b_0 + b_1 X - b_2 W + b_3 T + u$ حيث $W = X^2$ ، $T = X^3$

٨ — ٢٢ المطلوب توفيق دالة لوغاريتمية مزدوجة لبيانات جدول (٦ — ١٢) .

الإجابة :

$$\underset{(14.69)}{\ln Y = 2.64} + \underset{(6.31)}{0.72 \ln X} \qquad R^2 = 83.26\%$$

٨ — ٢٣ المطلوب توفيق دالة نصف لوغاريتمية على الصورة $Y = b_0 + b_1 \ln X + u$ لبيانات جدول (٦ — ١٢)

الإجابة :

$$Y = \underset{(0.36)}{2.62} + \underset{(5.90)}{27.12 \ln X} \qquad R^2 = 81.29\%$$

٨ — ٢٤ (أ) المطلوب توفيق كثيرة الحدود على الصورة $Y = b_0 + b_1 X - b_2 X^2 + u$ لبيانات جدول (٦—١٢) .

(ب) أيها يعطى نتائج أفضل لبيانات جدول (٦—١٢)، الصورة الخطية للمسائل (٦—٣٤) ، (٦ — ٣٧) ، (٦—٣٨) ، و(٦ — ٤٠) ، أم الصورة نصف اللوغاريتمية للمسألة (٨ — ٢٣) أم الصورة كثيرة الحدود الواردة فى (أ)

الإجابة : (أ)

$$Y = -2.25 + \underset{(1.99)}{13.67X} - \underset{(-1.14)}{0.77X^2} \qquad R^2 = 80.75\% \quad F_{2,7} = 14.68$$

(ب) التوفيق نصف اللوغاريتمى أفضل من الخطى وكثيرة الحدود

المتغيرات الصورية :

٨ — ٢٥ بالنسبة لبيانات جدول (٨ — ٢) (أ) قدر معادلة انحدار (٨ — ٦) (ب) هل يختلف معامل الميل جوهرياً فى وقت الحرب عن وقت السلم ؟ (جـ) ماهو معامل الميل فى وقت السلم ؟ فى وقت الحرب ؟

الإجابة : (أ)

$$\hat{Y} = -2.89 + \underset{(11.88)}{0.17X} - \underset{(-7.56)}{0.11XD} \qquad R^2 = 0.95$$

(ب) نعم (جـ) $\hat{b}_1 = 0.17$ فى وقت السلم و $\hat{b}_1 = 0.06$ فى وقت الحرب

٨ — ٢٦ بالنسبة لبيانات جدول (٨ — ٢) (أ) قدر معادلة انحدار (٨ — ٧) (ب) هل يختلف المقطع معنوياً فى وقت الحرب عن وقت السلم ؟ (جـ) هل يختلف معامل الميل معنوياً فى وقت الحرب عن وقت السلم ؟

الإجابة : (أ)

$$\hat{Y} = -3.34 + \underset{(11.58)}{0.17X} + \underset{(0.67)}{14.59D} - \underset{(-1.64)}{0.18XD} \qquad R^2 = 0.95$$

(ب) لا (جـ) لا

٨ — ٢٧ يعطى جدول (٨ — ١٥) صافى المبيعات لصناعات السلع المعمرة فى الولايات المتحدة ، S ، من الربع الأول لعام ١٩٧٤ حتى الربع الثالث لعام ١٩٧٩ . (أ) اختبر وجود اتجاه عام خطى فى نمو المبيعات ووجود تأثيرات موسمية . (ب) ماقيمة المقطع لكل موسم ؟

جدول (٨ — ١٥) صافى مبيعات صناعات السلع المعمرة (ببلايين الدولارات)

السنة	الربع: الأول	الثانى	الثالث	الرابع
1974	120.3	136.8	134.8	137.1
1975	121.3	132.4	131.0	136.3
1976	137.8	153.7	146.2	151.8
1977	151.2	169.5	163.8	172.7
1978	170.1	195.0	189.7	205.9
1979	208.1	223.3	214.6	

المصدر : التقرير الاقتصادى للرئيس ، مكتب حكومة الولايات المتحدة للطباعة ، واشنطن ، ١٩٨٠ صفحة ٣٠٠

الإجابة : (أ) بتعيين قيمة للاتجاه العام ، T ، مساوية 23,....,3,2,1 على التتابع لكل ربع سنة وبوضع $D_1 = 1$ للربع الثانى و 0 لغير ذلك ، $D_2 = 1$ للربع الثالث و 0 لغير ذلك ، $D_3 = 1$ للربع الرابع و 0 لغير ذلك ، نحصل على

$$\hat{S} = \underset{(13.69)}{103.60} + \underset{(2.17)}{4.35T} + \underset{(0.54)}{12.63D_1} + \underset{(0.81)}{3.18D_2} + 4.94D_3 \qquad R^2 = 0.92$$

(ب) حيث أن D_1 فقط معنوية إحصائياً عند مستوى 5% ، $\hat{b}_0 = 103.60$ فى ربع السنة الأول والثالث والرابع ، بينما $\hat{b}_0 = 116.23$ فى ربع السنة الثانى

٨ — ٢٨ يعطى جدول (٨ — ١٦) متوسط الدخل الأسبوعى ، Y ، بالدولار لعمال الإنتاج فى الصناعة ونسبة خريجى المدارس الثانوية للسكان سن 18 سنة فأكثر ، X لشرق الولايات المتحدة فى ١٩٧٦ . (أ) أجرانحدار Y على X وعلى متغيرات صورية تأخذ التأثيرات الموسمية فى الحسبان . (ب) ماقيمة المقطع لكل منطقة ؟

جدول (٨ — ١٦) الدخل ونسبة خريجى المدارس الثانوية فى الشرق عام ١٩٧٦

الدخل	166	168	180	190	164	209	208	216	210
نسبة الحاصلين على الثانوية العامة	67.8	70.3	69.7	72.3	61.7	70.3	66.2	66.4	64.8
الولاية	Maine	N.H.	Vt.	Mass.	R.I.	Conn.	N.Y.	N.J.	Pa.
المنطقة	نيوانجلند						وسط الأطلنطى		

الدخل	220	219	172	212	149	158	164	176
نسبة الحاصلين على الثانوية العامة	69.5	69.3	64.2	53.3	55.3	57.1	58.7	64.8
الولاية	Del.	Md.	Va.	W.Va.	N.C.	S.C.	Ga.	Fla.
المنطقة	جنوب الأطلنطى							

المصدر : التقرير الاقتصادى للرئيس ، مكتب حكومة الولايات المتحدة للطباعة ، واشنطن ، ١٩٨٠ صفحة ٤٢٤ ، ١٤٦ .

الإجابة : (أ) بأخذ نيوانجلند كأساس ، $D_1 = 1$ لولايات « وسط الأطلنطى » ، و 0 لغير ذلك ، $D_2 = 1$ لولايات « الأطلنطى » و 0 لغير ذلك ، نحصل على

$$\hat{Y} = 10.80 + \underset{(2.26)}{2.46X} + \underset{(2.66)}{38.92D_1} + \underset{(1.63)}{21.83D_2} \qquad R^2 = 0.45$$

(ب) $\hat{b}_0 = 10.80$ لولايات نيوانجلند وجنوب الأطلنطى ، بينما $\hat{b}_0 = 49.82$ لولايات وسط الأطلنطى

نماذج فترات الإبطاء الموزعة :

٨ – ٢٩ ماهى المشاكل فى تقرير : (أ) معادلة (٨ – ٩) ؟ (ب) معادلة (٨ – ١٠) ؟ (جـ) معادلة (٨ – ١٣)) ؟

الإجابة : (أ) تضييع مشاهدة مقابل كل قيمة مبطأة من X وعلى الأرجح سوف يكون هناك ارتباط بين قيم X بعضها البعض .
(ب) جمود هيكل الإبطاء على صورة تناقص هندسى وخرق اثنين من فروض OLD مما يؤدى إلى مقدرات متحيزة وغير متسقة .
(جـ) عدد المعاملات المطلوب تقديرها لايقل بنفس الدرجة كما فى معادلة (٨ – ١٠) وقد لايكون معروفاً طول وشكل فترة الإبطاء .

٨ – ٣٠ يعطى جدول (٨ – ١٧) إنفاق قطاع الأعمال على المعدات الجديدة للمرافق العامة Y ، والدخل القومى الإجمالى X ، كليهما ببلايين الدولارات ، للولايات المتحدة من ١٩٦٠ إلى ١٩٧٩ . (أ) قدر نموذج كويك (أى معادلة (٨ – ١٠)) . (ب) ماقيمة كل من $\hat{\lambda}$ و $\hat{a}$ ؟

الإجابة : (أ)

$$\hat{Y}_t = -1.92 + \underset{(4.55)}{0.01X_t} + \underset{(2.63)}{0.40Y_{t-1}} \qquad R^2 = 0.99$$

(ب) $\hat{\lambda} = 0.40$ و $\hat{a} = -320$.

جدول (٨ – ١٧) إنفاق قطاع الأعمال على المعدات الجديدة للمرافق العامة والدخل القومى الإجمالى : الولايات المتحدة ١٩٦٠ – ١٩٧٩ (ببلايين الدولارات)

السنة	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969
Y	5.2	5.0	4.9	5.0	5.5	6.3	7.4	8.7	10.2	11.6
X	506.0	523.3	563.8	594.7	635.7	688.1	753.0	796.3	868.5	935.5
السنة	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
Y	13.1	15.3	17.0	18.7	20.6	20.1	22.3	25.8	29.5	33.2
X	982.4	1063.4	1171.1	1306.6	1412.9	1528.8	1702.2	1899.5	2127.6	2368.5

المصدر : التقرير الاقتصادى للرئيس ، مكتب حكومة الولايات المتحدة للطباعة واشنطن ، ١٩٨٠ صفحة ٢٠٣ ، ٢٥٥

٨ – ٣١ يعطى جدول (٨ – ١٨) إجمالى الإنفاق الاستهلاكى الشخصى ، Y ، وإجمالى الدخل المتاح الشخصى ، X ، كليهما ببلايين الدولارات ، للولايات المتحدة من ١٩٦٠ إلى ١٩٧٩ . (أ) قدر نموذج ألمون للإبطاء بافتراض إبطاء لثلاث فترات على شكل كثيرة حدود من الدرجة الثانية (ب) هل يناسب هذا النموذج البيانات جيداً ؟

جدول (٨ – ١٨) الاستهلاك والدخل المتاح (ببلايين الدولارات) : الولايات المتحدة ١٩٦٠ – ١٩٧٩

السنة	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969
Y	324.9	335.9	355.2	374.6	400.4	430.2	464.8	490.4	535.9	579.7
X	349.4	362.9	383.9	402.8	437.0	472.2	510.4	544.5	588.1	630.4
السنة	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
Y	618.8	668.2	733.0	809.9	889.6	979.1	1089.9	1210.0	1350.8	1509.8
X	685.9	742.8	801.3	901.7	984.6	1086.7	1184.5	1305.1	1458.4	1623.2

المصدر : التقرير الاقتصادى للرئيس ، مكتب الحكومة الأمريكية للطباعة ، واشنطن ١٩٨٠ صفحة ٢٢٩ .

الإجابة : (أ) $R^2 = 0.99$ $\hat{Y} = -19.08 + 1.94X_t + 0.77X_{t-1} + 0.14X_{t-2} + 0.04X_{t-3}$
$(0.98) \quad (2.62) \quad (0.36) \quad (0.13)$

(ب) حيث أن معامل X_{t-1} فقط (أى $\hat{b}_1$) معنوى إحصائياً عند مستوى 5% وتتجاوز قيمته قيمة $\hat{b}_0$ ، فإن هذا النموذج لايناسب البيانات جيداً . قد يكون نموذج كويك أو صورة أخرى من نموذج ألمون أكثر ملاءمة .

التنبؤ :

٨ – ٣٢ عندما $X = 4$ فى المسألة (٦ – ٤٤) ، أوجد (أ) s_F^2 ، (ب) $\hat{Y}_F$ ، و (جـ) فترة الثقة 95% للتنبؤ Y_F .

الإجابة : (أ) $s_F^2 \cong 1.19$ (ب) $\hat{Y}_F = 4.78$ (جـ) 4.78 ± 2.43

٨ – ٣٣ فى المسألة (٧ – ٢٠) وعند $X_{1F} = 2$ و $X_{2F} = 1,250$ لعام ١٩٨٦ ، (أ) أوجد s_F^2 (ب) أوجد فترة الثقة 95% للتنبؤ Y_F ، بمعلومية أن

for Y_F, given that $\hat{Y} = 82.27 - 5.11X_1 + 0.02X_2$, $\bar{X}_1 = 6$, $\bar{X}_2 = 1{,}100$, $s^2 = \sum e^2/n - k = 226.32/12 \cong 18.86$, $s_{\hat{b}_1}^2 \cong 1.41$, $s_{\hat{b}_2}^2 \cong 0.01$, $s_{\hat{b}_0}^2 \cong 238.19$, and $s_{\hat{b}_1\hat{b}_2} \cong 0.01$.

الإجابة : (أ) $s_F^2 \cong 468.61$ (ب) $97.05 \pm (2.18)(21.65)$ أى بين 48.85 و 144.25 .

الفصل التاسع

مشاكل فى تحليل الانحدار

٩-١ تعدد العلاقات الخطية

يشير تعدد العلاقات الخطية إلى الحالة التى يكون فيها بين اثنين أو أكثر من المتغيرات المفسرة فى نموذج الانحدار ارتباط قوى ، مما يجعل من الصعب أو المستحيل عزل تأثيراتها الفردية على المتغير التابع . فى وجود هذا التعدد ، فإن معاملات OLS المقدرة قد تكون غير معنوية إحصائياً (وقد تأخذ الإشارة الخطأ) بالرغم من أن R^2 قد تكون « عالية » . وأحياناً يمكن التغلب على تعدد العلاقات الخطية أو اختزاله بجمع بيانات أكثر ، وباستخدام معلومات مسبقة ، بتحويل العلاقة الدالية (أنظر مسألة ٩ - ٣) ، أو بالتخلص من واحد من المتغيرات ذات الارتباط العالى .

مثال ١ : يعطى جدول ٩ - ١ مستوى الواردات ، Y والناتج القومى الإجمالى X_1 ، GNP كليهما ببلايين الدولارات ، والرقم القياسى العام لأسعار المستهلكين X_2 ، للولايات المتحدة من ١٩٦٤ إلى ١٩٧٩ . ومن المتوقع أن يرتفع مستوى الواردات مع زيادة GNP وارتفاع مستوى الأسعار المحلية . باجراء انحدار Y على X_1 و X_2 ، نحصل على

$$\hat{Y} = -101.49 + 0.08X_1 + 0.76X_2 \qquad R^2 = 0.97$$
$$\qquad (1.40) \qquad (1.00) \qquad \bar{R}^2 = 0.985$$
$$r_{12} = 0.997$$

وحيث أن كلا من $\hat{b}_1$ ، $\hat{b}_2$ ليست معنوية إحصائياً عند مستوى معنوية 5% بينما $R^2 \cong 0.99$ ، فإن هناك إشارة واضحة لوجود ازدواج خطى بين X_1 و X_2 . ويؤكد هذا معامل الارتباط البسيط المرتفع جداً بين X_1 و X_2 $(r_{12} \cong 1.0)$ باعادة تقدير الانحدار بدون X_1 أو X_2 نحصل على

$$\hat{Y} = -69.03 + 0.13X_1 \qquad R^2 = 0.986$$
$$(-12.00) \qquad (31.87)$$

$$\hat{Y} = -146.52 + 1.82X_2 \qquad R^2 = 0.985$$
$$(-17.58) \qquad (30.79)$$

فى الانحدار البسيط ، كل من C_1 ، X_2 معنوى إحصائياً عند مستوى أفضل من 1% . ولكن استبعاد أى منهما من علاقة الانحدار يؤدى إلى تقديرات OLS متحيزة ، لأن النظرية الاقتصادية تشير إلى وجوب دخول كل من GNP ومستوى الأسعار فى دالة الواردات .

جدول ٩ - ١ الواردات و GNP (كلاهما بالبليون دولار) والرقم القياسى لأسعار المستهلكين : الولايات المتحدة ١٩٦٤-١٩٧٩

السنة	1964	1965	1966	1967	1968	1969	1970	1971
Y	28.4	32.0	37.7	40.6	47.7	52.9	58.5	64.0
X_1	635.7	688.1	753.0	796.3	868.5	935.5	982.4	1,063.4
X_2	92.9	94.5	97.2	100.0	104.2	109.8	116.3	121.3
السنة	1972	1973	1974	1975	1976	1977	1978	1979
Y	75.9	94.4	131.9	126.9	155.4	185.8	217.5	260.9
X_1	1,171.1	1,306.6	1,412.9	1,528.8	1,702.2	1,899.5	2,127.6	2,368.5
X_2	125.3	133.1	147.7	161.2	170.5	181.5	195.4	217.4

المصدر : التقرير الاقتصادى للرئيس : مكتب حكومة الولايات المتحدة للطباعة واشنطن ، ١٩٨٠ ، صفحة ٢٠٣ ، ٢٥٩ .

٩–٢ اختلاف التباين

إذا لم يتوفر شرط OLS أن تباين حد الخطأ ثابت بالنسبة لكل قيم المتغيرات المستقلة ، فإننا نواجه مشكلة **اختلاف التباين** . ويؤدى هذا إلى تقديرات متحيزة وغير كفؤ (أى أكبر من أصغر تباين) للأخطاء المعيارية (وبالتالى اختبارات إحصائية وفترات ثقة خاطئة) . وأحد اختبارات الكشف عن اختلاف التباين يتضمن ترتيب البيانات من القيم الأصغر إلى القيم الأكبر للمتغير المستقل X وإجراء انحدارين ، واحد للقيم الصغيرة للمتغير X والآخر للقيم الكبيرة ، مع حذف ، خمس المشاهدات الوسطية مثلا . ثم نختبر ما إذا كانت نسبة مجموع مربعات الخطأ (ESS) للانحدار الثانى إلى الانحدار الأول تختلف معنوياً عن الواحد ، باستخدام جداول F بدرجات حرية $(n-d-2k)/2$ ، حيث n إجمالى عدد المشاهدات ، d عدد المشاهدات المحذوفة ، k عدد المعالم المقدرة .

أما إذا كان تباين الخطأ يتناسب مع X^2 (وهذا غالباً ما يحدث) ، فإنه يمكن التغلب على اختلاف التباين بقسمة كل حدود النموذج على X ثم إعادة تقدير الانحدار باستخدام المتغيرات المحولة .

مثال ٢ : يعطى جدول ٩ – ٢ متوسط الأجور Y ، وعدد العاملين X ، فى 30 شركة فى إحدى الصناعات . باجراء انحدار Y على X للعينة كلها ، نحصل على

$$\hat{Y} = \underset{(40.27)}{7.5} + \underset{(16.10)}{0.009X} \qquad R^2 = 0.90$$

نتائج انحدار Y على X للإثنى عشرة مشاهدة الأولى والإثنى عشرة مشاهدة الأخيرة هى ، على الترتيب

$$\hat{Y} = \underset{(39.4)}{8.1} + \underset{(4.36)}{0.006X} \qquad R^2 = 0.66 \quad ESS_1 = 0.507$$

$$\hat{Y} = \underset{(4.16)}{6.1} + \underset{(3.89)}{0.013X} \qquad R^2 = 0.60 \quad ESS_2 = 3.095$$

وحيث أن $ESS_2/ESS_1 = 3.9095 / 0.507 = 6.10$ تتجاوز $F_{10,10} = 2.97$ عند مستوى معنوية 5% (أنظر ملحق ٧) ، نقبل فرض اختلاف التباين . وباعادة تقدير النموذج المحول لتصحيح اختلاف التباين ، نحصل على

$$\hat{Y}/X = \underset{(14.43)}{0.008} + \underset{(76.58)}{7.8(1/X)} \qquad R^2 = 0.99$$

لاحظ أن معامل الميل يمثله الآن المقطع (أى 0.008) ، وأنه أصغر الآن من ذى قبل (أى 0.009) .

جدول ٩ – ٢ متوسط الأجور وعدد العاملين

عدد العاملين	متوسط الأجور					
100	9.00	8.90	8.70	8.60	8.40	8.40
200	9.60	9.40	9.30	9.30	9.10	8.90
300	10.50	10.30	10.30	9.90	9.80	9.50
400	11.70	11.50	11.30	10.90	10.60	10.30
500	13.10	12.70	12.50	12.10	11.80	11.60

٩–٣ الارتباط الذاتى

عندما يكون حد الخطأ فى فترة زمنية مرتبطاً طردياً مع حد الخطأ فى الفترة الزمنية السابقة عليها ، فإننا نواجه مشكلة **الارتباط الذاتى** (موجب ومن الدرجة الأولى) . وهذا شائع فى تحليل السلاسل الزمنية ويؤدى إلى أخطاء معيارية متحيزة إلى أسفل (وبالتالى إلى اختبارات إحصائية وفترات ثقة خاطئة) .

ويختبر وجود ارتباط ذاتى من الدرجة الأولى باستخدام جدول إحصائية ديربين – واتسون (ملحق ٨) عند مستويات معنوية 5% أو 1% لعدد n مشاهدات و k' متغيرات مفسرة . فإذا كانت القيمة d المحسوبة باستخدام معادلة (٩ – ١) أصغر من القيمة الجدولية d_L (الحد الأدنى) ، نقبل فرض وجود ارتباط ذاتى موجب من الدرجة الأولى .

$$d = \frac{\sum_{t=2}^{n} (e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2} \qquad (٩ - ١)$$

ويرفض الفرض فى حالة $d < d_U$ (الحد الأعلى) ، ويكون الاختبار غير حاسم فى حالة $d_L < d < d_U$ (وبالنسبة للارتباط الذاتى السالب ، أنظر المسألة ٩ – ٨) .

وكطريقة لتصحيح النموذج لوجود ارتباط ذاتى نقدر أولا ρ (الحرف اليونانى رو) من المعادلة (٩ – ٢) :

$$Y_t = b_0(1 - \rho) + \rho Y_{t-1} + b_1 X_t - b_1 \rho X_{t-1} + v_t \qquad (٩ - ٢)$$

ثم يعاد تقدير الانحدار على المتغيرات المحولة :

$$(Y_t - \hat{\rho} Y_{t-1}) = b_0(1 - \hat{\rho}) + b_1(X_t - \hat{\rho} X_{t-1}) + (u_t - \hat{\rho} u_{t-1}) \qquad (٩ - ٣)$$

ولتجنب ضياع المشاهدة الأولى فى عملية إيجاد الفروق ، نستخدم $Y_1\sqrt{1 - \hat{\rho}^2}$ و $X_1\sqrt{1 - \hat{\rho}^2}$ المشاهدة الأولى المحولة لكل من Y و X على الترتيب . وعندما تكون $\hat{\rho} \cong 1$ يمكن تصحيح الارتباط الذاتى باعادة إجراء الانحدار على شكل فروق وحذف حد المقطع (أنظر المسألة ٩ – ١٢) .

مثال ٣ – يعطى جدول ٩ – ٣ مستوى المخزون ، Y والمبيعات X ، كليهما بالبليون دولار ، فى الصناعة التحويلية الأمريكية من ١٩٥٩ إلى ١٩٧٨ ، باجراء انحدار Y على X ، نحصل على

$$\hat{Y}_t = \underset{(1.98)}{6.61} + \underset{(32.00)}{1.63 X_t} \qquad R^2 = 0.98 \quad d = 0.70$$

وحيث أن $d = 0.70 < d_L = 1.20$ عند مستوى معنوية 5% مع $n = 20$ و $k' = 1$ (من ملحق ٨) ، فهناك دليل على وجود ارتباط ذاتى . ويعطى معامل ρ فى الانحدار التالى تقديراً للمعامل Y_{t-1} :

$$\hat{Y}_t = 4.08 + \underset{(2.85)}{0.74 Y_{t-1}} + \underset{(3.10)}{1.49 X_t} - \underset{(-1.30)}{1.11 X_{t-1}} \qquad R^2 = 0.99$$

باستخدام $\hat{\rho} = 0.74$ لتحويل المتغيرات الأصلية ، كما فى معادلة (٩ – ٣) ، وباستخدام $52.9\sqrt{1 - 0.74^2} = 35.58$ و $30.3\sqrt{1 - 0.74^2} = 20.38$ المشاهدة المحولة الأولى لكل من Y و X على الترتيب ، نعيد إجراء الانحدار على المتغيرات المحولة (المميزة بالنجمة) ونحصل على

$$Y_t^* = \underset{(1.77)}{4.14} + \underset{(13.99)}{1.49 X_t^*} \qquad R^2 = 0.92 \quad d = 1.46$$

وحيث أنه الآن $d = 1.46 > d_U = 1.41$ (من ملحق ٨) ، فليس هناك دليل على وجود الارتباط الذاتى . لاحظ أن قيمة t للمتغير X_t^* أقل منها بالنسبة للمتغير X_t (ولكنها لازالت عالية المعنوية) وأن R^2 أيضاً أقل .

جدول ٩ – ٣ المخزون والمبيعات (بالبليون دولار) فى الصناعة التحويلية الأمريكية ، ١٩٥٩ – ١٩٧٨

السنة	1959	1960	1961	1962	1963	1964	1965	1966	1967	1968
Y	52.9	53.8	54.9	58.2	60.0	63.4	68.2	78.0	84.7	90.6
X	30.3	30.9	30.9	33.4	35.1	37.3	41.0	44.9	46.5	50.3
السنة	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978
Y	98.2	101.7	102.7	108.3	124.7	157.9	158.2	170.2	180.0	198.0
X	53.5	52.8	55.9	63.0	73.0	84.8	86.6	98.8	110.8	124.7

المصدر : التقرير الاقتصادى للرئيس . مكتب حكومة الولايات المتحدة للطباعة ، واشنطن ، ١٩٨٠ صفحة ٢٥٦ .

٩ـ٤ اخطاء فى المتغيرات

تشير الأخطاء فى المتغيرات إلى الحالة التى تحتوى فيها متغيرات الانحدار على أخطاء فى القياس . إن أخطاء القياس فى المتغير التابع تدخل فى حد التشويش ولا تخلق أى مشكلة . ولكن الأخطاء فى المتغيرات المفسرة تؤدى إلى تقديرات للمعالم متحيزة وغير متسقة .

وإحدى الطرق للحصول على تقديرات معالم OLS متسقة هى أن نستخدم ، بدلا من المتغير المفسر المتضمن أخطاء فى القياس ، متغيراً آخر (يسمى متغيراً وسيطاً) يكون ذا ارتباط عال مع المتغير المفسر الأصلى ولكنه مستقل عن حد الخطأ . وغالباً ما يكون القيام بهذا صعباً ويتضمن قدراً من التحكم . وأبسط متغير وسيط هو استخدام المتغير المفسر المبطأ (أنظر مثال ٤) . وطريقة أخرى تستخدم عندما تتضمن X وحدها أخطاء فى القياس وتتلخص فى إيجاد انحدار Xعلى Y (المربعات الصغرى المعكوسة ، أنظر المسألة ٩ – ١٥) .

مثال ٤ : يعطى جدول ٩ – ٤ المخزون Y ، والمبيعات الفعلية X ، وقيم افتراضية للمتغير X تشمل أخطاء فى القياس X' ، كلها بالبليون دولار ، لتجارة التجزئة الأمريكية من ١٩٦٣ إلى ١٩٧٨ . ويفترض أن X و Y خاليتان من الخطأ باجراء انحدار Y_tعلى X_t نحصل على

$$\hat{Y}_t = -1.92 + 1.53X_t \qquad R^2 = 0.996$$
$$(-1.79) \qquad (56.34) \qquad d = 1.86$$

باجراء انحدار Y_t على X'_t (فى حالة عدم توفر X_t) ، نحصل على

$$\hat{Y}_t = 0.74 + 1.32X'_t \qquad R^2 = 0.996$$
$$(0.73) \qquad (57.01) \qquad d = 1.88$$

لاحظ أن $b'_0 > b_0$ ، بينما $b'_1 < b_1$. باستخدام قيم X'_{t-1} كمتغير وسيط بدلا من X'_t (إذا كان هناك شك بأنَّ X'_t مرتبطة مع u_t) ، نحصل على

$$\hat{Y}'_t = -1.56 + 1.50X'_{t-1} \qquad R^2 = 0.993$$
$$(-1.13) \qquad (44.47) \qquad d = 2.19$$
$$r_{X'_tX'_{t-1}} = 0.998$$

إن تقديرات المعالم الجديدة أقرب الآن إلى المعالم الحقيقية ومتسقة .

جدول ٩ – ٤ المخزون والمبيعات (بالبليون دولار) فى تجارة التجزئة الأمريكية : ١٩٦٣ – ١٩٧٨

السنة	1963	1964	1965	1966	1967	1968	1969	1970
Y	29.4	31.1	34.4	38.1	35.3	38.9	42.5	43.9
X	20.6	21.8	23.7	25.3	24.4	27.0	28.9	30.7
X'	21.9	23.3	25.5	27.3	26.3	29.3	31.5	33.6
السنة	1971	1972	1973	1974	1975	1976	1977	1978
Y	50.1	55.1	63.2	71.1	71.1	79.3	90.1	100.8
X	33.9	37.4	41.9	44.7	48.7	54.6	60.3	66.6
X'	37.3	41.4	46.6	49.9	54.5	61.3	67.9	75.3

المصدر : التقرير الاقتصادى للرئيس . مكتب الحكومة الأمريكية للطباعة واشنطن ١٩٨٠ ، صفحة ٢٥٦ .

مسائل محلولة

تعدد العلاقات الخطية :

٩ – ١ (أ) ماذا يقصد بتعدد العلاقات الخطية التام ؟ ما هو تأثيره ؟ (ب) ماذا يقصد بتعدد مرتفع ولكن ليس تاماً ؟ ما هى المشاكل التى يمكن أن تنشأ عنه ؟ (جـ) كيف يمكن اكتشاف تعدد العلاقات الخطية ؟
(د) ماذا يمكن عمله للتغلب على أو اختزال المشاكل الناجمة عن تعدد العلاقات الخطية ؟

(أ) يكون بين متغيرين مستقلين أو أكثر تعدد خطى تام إذا كان من الممكن التعبير عن واحد أو أكثر من المتغيرات كمزيج خطى للمتغير (المتغيرات) الآخر. فمثلا يكون هناك تعدد خطى تام بين X_1 و X_2 إذا كانت $X_1 = 2X_2$ أو $X_1 = 5 - \frac{1}{3}X_2$ إذا كان هناك تعدد خطى تام بين متغيرين مستقلين أو أكثر ، سيكون من غير الممكن تقدير معالم OLS لأن مجموعة المعادلات الطبيعية سوف يكون بينها معادلة أو أكثر غير مستقلة .

(ب) يشير تعدد العلاقات الخطية المرتفع ، ولكن غير التام ، إلى الحالة التى يكون فيها بين متغيرين مستقلين أو أكثر فى النموذج الخطى ارتباط مرتفع . ويجعل هذا من الصعب أو من غير الممكن عزل تأثير كل متغير مفسر ، من بين المتغيرات ذات الارتباط المرتفع فيما بينها ، على المتغير التابع . ولكن ، تظل تقديرات OLS غير متحيزة (إذا كان النموذج قد حدد بدقة) . بالإضافة إلى أنه إذا كان الهدف الرئيسى هو التنبؤ ، لا يمثل الازدواج الخطى مشكلة إذا استمر أيضاً نمط التعدد خلال فترة التنبؤ .

(جـ) الحالة الكلاسيكية للتعدد الخطى تحدث عندما لا يكون أى من المتغيرات المفسرة فى انحدار OLS معنوياً إحصائياً (وبعضها قد يأخذ الإشارة الخطأ)، بالرغم من أن R^2 قد تكون عالية (مثلا ، بين 0.7 و 1.0) . فى الحالات الأقل وضوحاً ، قد يكون اكتشاف التعدد الخطى أكثر صعوبة . أحيانا يستخدم المعامل المرتفع للارتباط البسيط أو للارتباط الجزئى بين المتغيرات المفسرة كمقياس للتعدد الخطى . ولكن قد يوجد تعدد خطى خطير ، حتى لو كان الارتباط البسيط أو الجزئى منخفضاً نسبياً (أقل من 0.5 مثلا) .

(د) يمكن أحياناً تصحيح التعدد الخطى الكبير من خلال (١) زيادة حجم بيانات العينة ، (٢) استخدام المعلومات المسبقة (مثلا ، قد نعرف من دراسة سابقة أن $b_2 = 0.25b_1$) (٣) تحويل العلاقة الدالية ، أو (٤) حذف أحد المتغيرات ذات الارتباط المرتفع مع غيرها من المتغيرات (ولكن ، قد يؤدى هذا إلى تحيز أو خطأ فى تحديد النموذج إذا كانت النظرية تخبرنا أن المتغير المحذوف يجب أن يكون فى النموذج) .

٩ – ٢ يعطى جدول ٩ – ٥ الإنتاج بالأطنان Q ، مدخلات العمل عامل – ساعة L ، ومدخلات رأس المال ماكينة – ساعة K ، لعدد 15 شركة فى إحدى الصناعات والمطلوب (أ) توفيق دالة إنتاج كوب – دوجلاس فى صورة $Q = b_0 L^{b_1} K^{b_2} e^u$ للبيانات وإيجاد $\bar{R}^2$ ومعامل الارتباط البسيط بين $\ln L$ و $\ln K$ (ب) إيجاد انحدار $\ln Q$ على $\ln L$ فقط . (جـ) إيجاد انحدار $\ln Q$ على $\ln K$ فقط . (د) ما هى النتائج التى يمكن التوصل إليها بصدد تعدد العلاقات الخطية .

جدول ٩ – ٥ الإنتاج ، مدخلات العمل ورأس المال فى 15 شركة فى صناعة ما

الشركة	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Q	2350	2470	2110	2560	2650	2240	2430	2530	2550	2450	2290	2160	2400	2490	2590
L	2334	2425	2230	2463	2565	2278	2380	2437	2446	2403	2301	2253	2367	2430	2470
K	1570	1850	1150	1940	2450	1340	1700	1860	1880	1790	1480	1240	1660	1850	2000

(أ) بتحويل البيانات إلى صورة اللوغاريتمات الطبيعية كما هو موضح فى جدول ٩ – ٦ ثم بإيجاد انحدار $\ln Q$ على $\ln L$ و $\ln K$ ، نحصل على

$$\ln Q = 0.50 + \underset{(1.07)}{0.76 \ln L} + \underset{(1.36)}{0.19 \ln K} \qquad R^2 = 0.969 \quad \bar{R}^2 = 0.964 \quad r_{\ln L \ln K} = 0.992$$

(ب)

$$\ln Q = \underset{(-7.74)}{-5.50} + \underset{(18.69)}{1.71 \ln L} \qquad R^2 = 0.964$$

(جـ)

$$\ln Q = \underset{(4.78)}{5.30} + \underset{(19.19)}{0.34 \ln K} \qquad R^2 = 0.966$$

(د) حيث أن كلا من $\hat{b}_1$ ، $\hat{b}_2$ فى (أ) غير معنوية إحصائياً عند مستوى معنوية 5% (بمعنى أن لهما أخطاء معيارية كبيرة بدون داع) ، بينما $R^2 = 0.97$ ، فإن هناك إشارة واضحة لوجود تعدد علاقات خطية . أى أن ، الشركات الكبيرة تميل إلى استخدام عمل أكثر ورأس مال أكبر من الشركات الصغيرة . ويؤكد هذا معامل الارتباط البسيط المرتفع ، 0.99 بين $\ln L$ و $\ln K$. وقد أعيد تقدير معادلات الانحدار البسيط فى (ب) و (جـ) باستخدام إما $\ln K$ أو $\ln K$ كمتغير مفسر وحيد . وفى حالات الانحدار البسيط فان كلا من $\ln L$ و $\ln K$ معنوية إحصائياً عند مستوى أفضل من 1% مع R^2 تتجاوز 0.96 . ولكن حذف $\ln K$ أو $\ln L$ من الانحدار المتعدد يؤدى إلى تقدير متحيز لميل OLS للمتغير الباقى فى العلاقة لأن النظرية الاقتصادية تفترض مقدماً أن كلا من العمل ورأس المال يجب أن يكون فى دالة الإنتاج .

جدول ٩ – ٦ الإنتاج ، مدخلات العمل ورأس المال فى صورتها الأصلية وفى الصورة اللوغاريتمية

الشركة	Q	L	K	$\ln Q$	$\ln L$	$\ln K$
1	2350	2334	1570	7.76217	7.75534	7.35883
2	2470	2425	1850	7.81197	7.79359	7.52294
3	2110	2230	1150	7.65444	7.70976	7.04752
4	2560	2463	1940	7.84776	7.80914	7.57044
5	2650	2565	2450	7.88231	7.84971	7.80384
6	2240	2278	1340	7.71423	7.73105	7.20042
7	2430	2380	1700	7.79565	7.77486	7.43838
8	2530	2437	1860	7.83597	7.79852	7.52833
9	2550	2446	1880	7.84385	7.80221	7.53903
10	2450	2403	1790	7.80384	7.78447	7.48997
11	2290	2301	1480	7.73631	7.74110	7.29980
12	2160	2253	1240	7.67786	7.72002	7.12287
13	2400	2367	1660	7.78322	7.76938	7.41457
14	2490	2430	1850	7.82004	7.79565	7.52294
15	2590	2470	2000	7.85941	7.81197	7.60090

٩ – ٣ كيف يمكن التغلب على صعوبة التعدد الخطى فى المسألة ٩ – ٢ إذا كان من المعلوم أن هذه الصناعة تخضع لثبات الغلة (أى أن $b_1 + b_1 = 1$) ؟

عند ثبات الغلة ، يمكن إعادة كتابة دالة إنتاج كوب – دوجلاس كما يلى :

$$Q = b_0 + L^{b_1} K^{1-b_1} e^u$$

وبالتعبير عن دالة الإنتاج فى صورة لوغاريتمية مزدوجة وإعادة ترتيب الحدود ، نحصل على :

$$\ln Q = \ln b_0 + b_1 \ln L + (1 - b_1) \ln K + u$$

$$\ln Q - \ln K = \ln b_0 + b_1(\ln L - \ln K) + u$$

بوضع $\ln Q^* = \ln Q - \ln K$ و $\ln L^* = \ln L - \ln K$ ثم بايجاد انحدار $\ln Q^*$ على $\ln L^*$ نحصل على

$$\ln Q^* = \underset{(9.26)}{0.07} + \underset{(39.81)}{0.83} \ln L^* \qquad R^2 = 0.992$$

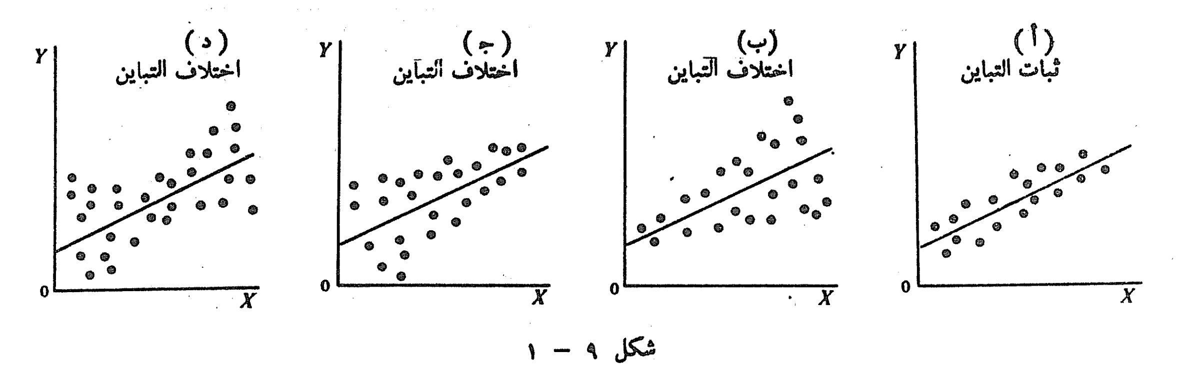
إذن $\hat{b}_2 = 1 - \hat{b}_1 = 1 - 0.83 = 0.17$.

اختلاف التباين :

٩ – ٤ (أ) ماذا يقصد باختلاف التباين ؟ (ب) ارسم شكلا يوضح عناصر تشويش لها تباين ثابت وكذلك الأشكال المختلفة لاختلاف التباين . (جـ) لماذا يمثل اختلاف التباين مشكلة ؟

(أ) يشير اختلاف التباين إلى الحالة التى يكون فيها تباين حد الخطأ غير ثابت عند كل قيم المتغير المستقل . أى أن ، $E(X_i u_i) \neq 0$ وعليه فإن $E(u_i)^2 \neq \sigma_u^2$ ويخرق هذا الفرض الثالث لنموذج انحدار OLS(أنظر المسألة ٦–٤) ويحدث هذا أساساً فى البيانات المقطعية . فمثلا ، تباين الخطأ الخاص بالإنفاق لعائلات الدخل المنخفض عادة يكون أصغر عنه بالنسبة لعائلات الدخل المرتفع لأن معظم إنفاق الأسر ذات الدخل المنخفض يكون على الضروريات ، مما يترك مجالات ضيقا لحرية الاختيار .

(ب) يوضح شكل ٩ – ١ (أ) حالة ثبات التباين لعناصر التشويش . بينما توضح أشكال ٩ – ١ ب ، جـ ، د ، اختلاف التباين لعناصر التشويش . فى شكل ٩ – ١ ب يزداد σ_u^2 مع X_i . فى شكل ٩ – ١ جـ يقل σ_u^2 مع X_i . فى شكل ٩ – ١ د يقل σ_u^2 ثم يزيد مع تزايد X_i . فى الاقتصاد ، اختلاف التباين كما فى ٩ – ١ ب هو الأكثر شيوعاً ، ومن ثم فإن المناقشة التالية تتعلق بهذه الحالة .



شكل ٩ – ١

(جـ) فى وجود حالة اختلاف التباين ، فإن تقديرات معالم OLS تظل غير متحيزة ومتسقة ولكنها تكون غير كفؤ (بمعنى أن لها تباينا أكبر من أقل تباين) . بالإضافة فإن تقديرات التباين تكون متحيزة ، مما يؤدى إلى اختبارات إحصائية غير صحيحة للمعالم وفترات ثقة متحيزة .

٩ – ٥ كيف يتم اختبار وجود حالة اختلاف التباين ؟ (ب) كيف يمكن تصحيح اختلاف التباين ؟

(أ) يمكن اختبار وجود حالة اختلاف التباين بترتيب البيانات من أصغر قيمة إلى أكبر قيمة من قيم المتغير المستقل ، X_i، وإجراء انحدارين منفصلين ، واحد للقيم الصغيرة ، وآخر للقيم الكبيرة للمتغير X_i ، مع حذف بعض المشاهدات الوسطية (خمس المشاهدات مثلا) . ثم يختبر نسبة مجموع مربعات الخطأ للانحدار الأول إلى مجموع مربعات الخطأ للانحدار الثانى (أى ESS_2 / ESS_1) لنرى هل تختلف معنوياً عن الواحد . ويستخدم توزيع F للاختبار

بدرجات حرية $(n - d - 2k)/2$ ، حيث n إجمالى عدد المشاهدات ، d عدد المشاهدات المحذوفة ، k عدد المعالم المقدرة . وهذا هو اختبار جولد فيلد – كوانت لاختلاف التباين وهو مناسب تماماً للعينات الكبيرة (أى عندما $n \geq 30$). إذا لم تحذف أية مشاهدات وسيطة ، يظل الاختبار صحيحاً ، ولكن قوته فى الكشف عن اختلاف التباين تكون أقل .

(ب) إذا افترض (وكثيراً ما يحدث هذا) أن $\text{var}\, u_i = CX_i^2$ ، حيث C ثابت يختلف عن الصفر ، فإنه يمكننا تصحيح اختلاف التباين بالقسمة (أى بترجيح) كل حد من حدود الانحدار على X_i ، ثم إعادة تقدير الانحدار باستخدام المتغيرات المحولة . فى حالة الانحدار ذى المتغيرين ، يكون لدينا

$$\frac{Y_i}{X_i} = \frac{b_0}{X_i} + b_1 + \frac{u_i}{X_i} \qquad (٩ - ٤)$$

ويصبح الآن حد الخطأ المحول ثابت التباين :

$$\text{var}\, u_i = \text{var}\,\frac{u_i}{X_i} = \frac{1}{X_i^2}\,\text{var}\, u_i = C\,\frac{X_i^2}{X_i^2} = C$$

لاحظ أن المقطع الأصلى أصبح متغيراً فى معادلة (٩-٤) بينما معلمة الميل الأصلية b_1 ، أصبحت هى المقطع الجديد . ولكن ، يجب توخى الحرص فى تفسير النتائج للانحدار المحول أو المرجح . حيث أن الأخطاء فى معادلة (٩ – ٤) ثابتة التباين ، ولذا فإن تقديرات OLS ليست فقط غير متحيزة ومتسقة ، ولكنها أيضاً كفء . وفى حالة الانحدار المتعدد ، يقسم كل حد فى الانحدار (أى يرجح) على المتغير المستقل (مثلا X_{2i}) الذى يظن أنه يرتبط مع حد الخطأ ، فيصبح لدينا

$$\frac{Y_i}{X_{2i}} = \frac{b_0}{X_{2i}} + b_1\frac{X_{1i}}{X_{2i}} + b_2 + \frac{u_i}{X_{2i}} \qquad (٩ - ٥)$$

فى معادلة ٩ – ٥ يصبح المقطع الأصلى b_0 ، متغيراً ، بينما يصبح b_2 المقطع الجديد . ويمكننا أن نحدد بالنظر إذا كانت X_{2i} أو X_{1i} هى المرتبطة مع u_i برسم كل من X_{2i} و X_{1i} مقابل بواقى الانحدار .

٩ – ٦ يعطى جدول ٩ – ٧ الانفاق الاستهلاكى C ، والدخل المتاح Y_d ، لعدد 30 أسرة (أ) أوجد انحدار C على Y_d للعينة ككل واختبر بالنسبة لاختلاف التباين . (ب) صحح بالنسبة لاختلاف التباين إن وجد فى (أ) .

جدول ٩ – ٧ بيانات الاستهلاك والدخل لعدد 30 أسرة

الاستهلاك			الدخل
\$10,600	\$10,800	\$11,100	\$12,000
11,400	11,700	12,100	13,000
12,300	12,600	13,200	14,000
13,000	13,300	13,600	15,000
13,800	14,000	14,200	16,000
14,400	14,900	15,300	17,000
15,000	15,700	16,400	18,000
15,900	16,500	16,900	19,000
16,900	17,500	18,100	20,000
17,200	17,800	18,500	21,000

(أ) باجراء انحدار C على Y_d للعينة كلها من 30 مشاهدة ، نحصل على :

$$\hat{C} = \underset{(3.29)}{1,480.0} + \underset{(29.37)}{0.788}\,Y_d \qquad R^2 = 0.97$$

للاختبار بالنسبة لاختلاف التباين ، نجرى انحدار C على Y_d لعدد 12مشاهدة الأولى ولعدد 12 مشاهدة الأخيرة ، مع حذف عدد ، مشاهدات الوسيطة ، ونحصل على

$$\hat{C} = \underset{(0.74)}{846.7} + \underset{(9.91)}{0.837\,Y_d} \qquad R^2 = 0.91 \qquad ESS_1 = 1{,}069{,}000$$

$$\hat{C} = \underset{(0.79)}{2{,}306.7} + \underset{(5.00)}{0.747\,Y_d} \qquad R^2 = 0.71 \qquad ESS_1 = 3{,}344{,}000$$

وحيث أن $ESS_2/ESS_1 = 3{,}344{,}000/1.096{,}000 = 3.13$ تتجاوز $F = 2.97$ بدرجات حرية $10 = 2/(4 - 6 - 30)$ فى البسط والمقام عند مستوى معنوية 5% (أنظر ملحق ٧) ، نقبل فرض وجود اختلاف التباين .

(ب) بافتراض أن تباين الخطأ يتناسب مع Y_d^2 ، وباعادة تقدير الانحدار باستخدام المتغيرات المحولة فى جدول ٩ – ٨ لتصحيح اختلاف التباين ، نحصل على فى العمود الأخير من ٩ – ٨ (0.833333E-04 = 0.0000833333) :

$$\frac{\hat{C}}{Y_d} = \underset{(31.51)}{0.792} + \underset{(3.59)}{1{,}421.3}\frac{1}{Y_d} \qquad R^2 = 0.32$$

لاحظ أن الميل الحدى للاستهلاك هو الآن المقطع (أى 0,792) وهو أكبر مما كان قبل التعديل (أى 0.788) . فإن المعنوية الإحصائية للمعلمتين المقدرتين أعلى الآن من ذى قبل . أما R^2 للانحدار المرجح (أى 0.32) فهى أقل كثيراً وإن كانت المقارنة المباشرة مع قيمة R^2 ، 0.97 ، قبل التحويل غير ممكنة ، لأن المتغيرات التابعة أصبحت مختلفة (Y/X مقابل Y) .

جدول ٩ – ٨ الاستهلاك والدخل فى الصورة الأصلية والمحولة

الأسرة	C	Y_d	C/Y_d	$1/Y_d$
1	10600	12000	0.883333	0.833333E-04
2	10800	12000	0.900000	0.833333E-04
3	11100	12000	0.925000	0.833333E-04
4	11400	13000	0.876923	0.769231E-04
5	11700	13000	0.900000	0.769231E-04
6	12100	13000	0.930769	0.769231E-04
7	12300	14000	0.878571	0.714286E-04
8	12600	14000	0.900000	0.714286E-04
9	13200	14000	0.942857	0.714286E-04
10	13000	15000	0.866667	0.666667E-04
11	13300	15000	0.886667	0.666667E-04
12	13600	15000	0.906667	0.666667E-04
13	13800	16000	0.862500	0.625000E-04
14	14000	16000	0.875000	0.625000E-04
15	14200	16000	0.887500	0.625000E-04
16	14400	17000	0.847059	0.588235E-04
17	14900	17000	0.876471	0.588235E-04
18	15300	17000	0.900000	0.588235E-04
19	15000	18000	0.833333	0.555556E-04
20	15700	18000	0.872222	0.555556E-04
21	16400	18000	0.911111	0.555556E-04
22	15900	19000	0.836842	0.526316E-04
23	16500	19000	0.868421	0.526316E-04
24	16900	19000	0.889474	0.526316E-04
25	16900	20000	0.845000	0.500000E-04
26	17500	20000	0.875000	0.500000E-04
27	18100	20000	0.905000	0.500000E-04
28	17200	21000	0.819048	0.476190E-04
29	17800	21000	0.847619	0.476190E-04
30	18500	21000	0.880952	0.476190E-04

٩ — ٧ يعطى جدول ٩ — ٩ مستوى المخزون I ، والمبيعات S ، كليهما بالمليون دولار ، ومعدلات الاقتراض لعدد 35 شركة فى إحدى الصناعات . ومن المتوقع أن I سوف تكون مرتبطة بطردياً مع S وعكسياً مع R (أ) أوجد انحدار I على S و R للعينة كلها واختبر وجود اختلاف التباين . (ب) صحح بالنسبة لإختلاف التباين إن وجد فى (أ) ، بافتراض أن تباين الخطأ يتناسب مع S^2

جدول ٩ — ٩ المخزون ، المبيعات ، ومعدلات الاقتراض لعدد 53 شركة

الشركة	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
I	10	10	10	11	11	11	12	12	12	12	12	13	13	13	14	14	14	15
S	100	101	103	105	106	106	108	109	111	111	112	113	114	114	116	117	118	120
R	17	17	17	16	16	16	15	15	14	14	14	14	13	13	12	12	12	11
الشركة	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	
I	15	15	15	16	16	16	17	17	17	17	18	18	19	19	19	20	20	
S	122	123	125	128	128	131	133	134	135	136	139	143	147	151	157	163	171	
R	11	11	11	10	10	10	10	9	9	9	8	8	8	8	8	7	7	

(أ) باجراء انحدار I على S و R للعينة كلها من 35 شركة ، نحصل على

$$\hat{I} = -6.17 + \underset{(12.39)}{0.20S} - \underset{(-2.67)}{0.25R} \qquad R^2 = 0.98$$

لاختبار اختلاف التباين ، نجرى انحدار I على S و R لعدد 14 مشاهدة الأولى ، ولعدد 14 مشاهدة الأخيرة ، مع حذف 7 مشاهدات وسطية ونحصل على

$$\hat{I} = -2.23 + \underset{(1.90)}{0.16S} - \underset{(-0.81)}{0.22R} \qquad R^2 = 0.94 \quad ESS_1 = 0.908$$

$$\hat{I} = 16.10 + \underset{(3.36)}{0.11S} - \underset{(-3.35)}{1.40R} \qquad R^2 = 0.96 \quad ESS_2 = 5.114$$

وحيث أن $ESS_2/ESS_1 = 5.114/0.908 = 5.63$ تتجاوز $F_{11,11} = 2.82$ عند مستوى معنوية 5% (أنظر ملحق ٧) ، فإننا نقبل فرض وجود اختلاف التباين .

(ب) بافتراض أن تباين الخطأ يتناسب مع S^2 وباعادة تقدير الانحدار باستخدام المتغير المحول لتصحيح اختلاف التباين ، نحصل على

$$\frac{\hat{I}}{S} = \underset{(12.34)}{0.21} - 8.45(1/S) - \underset{(-2.98)}{0.18(R/S)} \qquad R^2 = 0.93$$

$b_0 = 0.21$ هى الآن معامل الميل المرتبط بالمتغير S (بدلا من 0.16 قبل التحويل) ، بينما $b_2 = -0.18$ هى معامل الانحدار المرتبط بالمتغير R (بدلا من 0.25 — قبل التحويل) . ويبقى كل من معاملى الميل ذا معنوية إحصائية عالية قبل وبعد التحويل ، وكذلك R^2 . الثابت الجديد هو 8.45 — بدلا من 6.17 — .

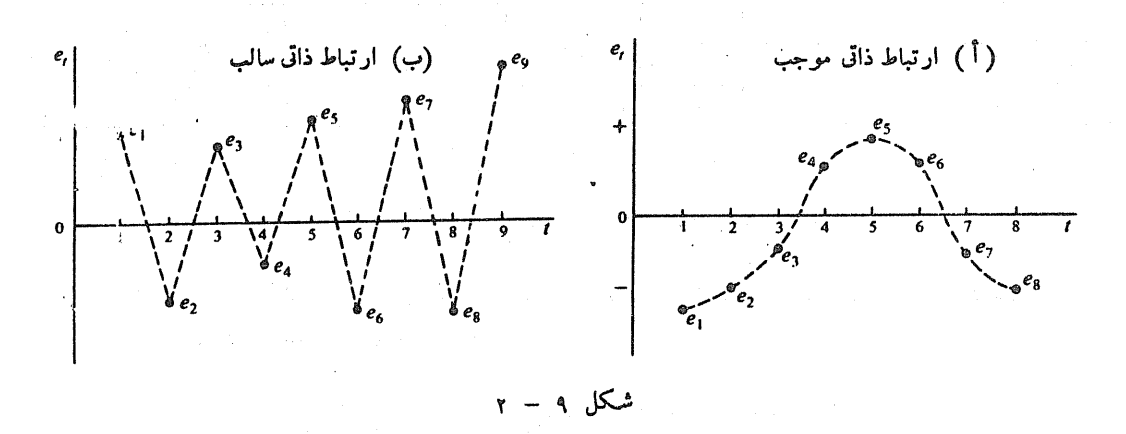
الارتباط الذاتى :

٩ — ٨ (أ) ماذا يقصد بالارتباط الذاتى ؟ (ب) ارسم شكلا يوضح ارتباطاً ذاتياً من الدرجة الأولى ، موجباً ، وسالباً (جـ) لماذا يعتبر الارتباط الذاتى مشكلة ؟

(أ) يشير **الارتباط الذاتى أو الارتباط المتسلسل** إلى الحالة التى يكون فيها حد الخطأ فى فترة زمنية على علاقة مع حد الخطأ فى أى فترة زمنية أخرى . إذا كان حد الخطأ فى فترة زمنية مرتبطاً بحد الخطأ فى الفترة الزمنية السابقة ، يكون هناك

ارتباط ذاتى من الدرجة الأولى . ومعظم التطبيقات فى الاقتصاد القياسى تتضمن ارتباطاً ذاتياً من الدرجة الأولى أكثر من الدرجة الثانية أو أكثر . وبالرغم من أنه من الممكن أن يكون هناك ارتباط ذاتى سالب ، فإن معظم السلاسل الزمنية الاقتصادية تظهر ارتباطا ذاتياً موجباً . ويعنى الارتباط الذاتى الموجب من الدرجة الأولى أن $E_{u_t u_{t-1}} > 0$ وفى هذا خرق لفرض OLS الرابع (أنظر مسألة ٦ — ٤) . وهذا شائع فى تحليل السلاسل الزمنية .

(ب) يوضح شكل ٩ — ٢ (أ) ارتباطاً ذاتياً موجباً من الدرجة الأولى ، بينما يوضح شكل ٩ — ٢ (ب) ارتباطاً ذاتياً سالباً من الدرجة الأولى . وعندما تكون لعدة بواقى متتالية نفس الإشارة كما فى شكل ٩ — ٢ (أ) ، يكون هناك ارتباط ذاتى موجب من الدرجة الأولى . ولكن عندما تغير البواقى المتتالية إشاراتها كثيراً ، كما فى شكل ٩ — ٢ (ب) يكون هناك ارتباط ذاتى سالب من الدرجة الأولى .



شكل ٩ — ٢

(جـ) فى وجود الارتباط الذاتى ، تظل تقديرات OLS غير متحيزة ومتسقة ، ولكن الخطأ المعيارى لمعالم الانحدار المقدرة تكون متحيزة ، مؤدية إلى اختبارات إحصائية غير صحيحة ، وإلى فترات ثقة متحيزة . وعندما يكون الارتباط الذاتى من الدرجة الأولى موجباً ، تكون الأخطاء المعيارية لمعالم الانحدار المقدرة متحيزة إلى أسفل ، ومن ثم يكون هناك مبالغة فى الدقة وفى المعنوية الإحصائية لمعالم الانحدار المقدرة .

٩ — ٩ (أ) كيف يمكن اختبار وجود الارتباط الذاتى من الدرجة الأولى الموجب أو السالب ؟ (ب) كيف يمكن تصحيح الارتباط الذاتى ؟

(أ) يمكن اختبار وجود الارتباط الذاتى بحساب إحصائية ديرين — واتسون ، d ، المعبر عنها بالمعادلة (٩ — ١) . وتعطى هذه الإحصائية بشكل روتينى كأحد نواتج معظم برامج الكمبيوتر مثل SPSS :

$$d = \frac{\sum_{t=2}^{n} (e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2} \qquad (٩ - ١)$$

(أنظر المسألة ٧ — ٢١) . وتتراوح القيمة المحسوبة d بين 0 و 4 ، ولا يكون هناك ارتباط ذاتى إذا كانت d قريبة من 2 . ويوضح شكل ٩ — ٣ قيم d التى تشير إلى وجود أو غياب ارتباط ذاتى من الدرجة الأولى موجب أو سالب ، أو التى تجعل الاختبار غير حاسم . وعندما يظهر المتغير التابع المبطأ كتغير مفسر فى الانحدار ، فإن d تكون متحيزة نحو 2 وتضعف قوتها فى الكشف عن الارتباط الذاتى .

| ارتباط ذاتى موجب | غير حاسم | لا يوجد ارتباط ذاتى | غير حاسم | ارتباط ذاتى سالب |

$0 \quad d_L \quad d_U \quad 2 \quad 4 - d_U \quad 4 - d_L \quad 4$

شكل ٩ – ٣

(ب) إن إحدى طرق تصحيح الارتباط الذاتى الموجب من الدرجة الأولى (النوع المعتاد) تتضمن أولا إجراء انحدار Y على قيمتها المبطأة لفترة واحدة ، وعلى متغيرات النموذج المفسرة ، وعلى المتغيرات المفسرة مبطأة لفترة واحدة .

$$Y_t = b_0(1 - \rho) + \rho Y_{t-1} + b_1 X_t - b_1 \rho X_{t-1} + v_t \qquad (٩ - ٢)$$

(وتشتق المعادلة السابقة بضرب كل من نموذج OLS الأصلى المبطأ لفترة زمنية واحدة فى ρ ، وطرح الناتج من نموذج OLS الأصلى ، مع نقل الحد ρY_{t-1} من الجانب الأيسر إلى الجانب الأيمن من المعادلة ، ووضع $v_t = u_t - \rho u_{t-1}$) . وتتضمن الخطوة الثانية استخدام قيمة ρ المقدرة فى معادلة (٩ – ٢) لتحويل كل متغيرات نموذج OLS الأصلى ، كما هو موضح فى معادلة (٩ – ٣) ، ثم إعادة تقدير معادلة (٩ – ٣) :

$$Y_t - \hat{\rho} Y_{t-1} = b_0(1 - \hat{\rho}) + b_1(X_t - \hat{\rho} X_{t-1}) + v_t \qquad (٩ - ٣)$$

ويصبح حد الخطأ الجديد ، v_t فى معادلة (٩ – ٣) خالياً من الارتباط الذاتى . ويعرف هذا الإجراء بطريقة ديربن على مرحلتين ويعتبر نموذجا للمربعات الصغرى العامة . ولتجنب فقدان المشاهدة الأولى فى عملية استخدام الفروق ، يستخدم $Y_1\sqrt{1 - \hat{\rho}^2}$ و $X_1\sqrt{1 - \hat{\rho}^2}$ للمشاهدة الأولى المحولة لكل من Y و X ، على الترتيب . أما إذا كان الارتباط الذاتى راجعاً إلى حذف متغير مهم ، أو شكل دالى خاطىء ، أو تحديد غير سليم للنموذج ، فيجب التخلص من هذه المشاكل أولا ، قبل تطبيق الإجراء السابق لتصحيح الارتباط الذاتى .

٩ – ١٠ يعطى جدول ٩ – ١٠ مستوى واردات الولايات المتحدة ، M و GNP (كليهما بالبليون دولار) من ١٩٦٠ إلى ١٩٧٩ (أ) أجر انحدار M على GNP واختبر بالنسبة لوجود الارتباط الذاتى عند مستوى معنوية 5 % (ب) صحح بالنسبة للارتباط الذاتى إن وجد فى (أ) .

جدول ٩ – ١٠ واردات الولايات المتحدة و GNP (بالبليون دولار) من ١٩٦٠ إلى ١٩٧٩

السنة	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969
M	23.2	23.1	25.2	26.4	28.4	32.0	37.7	40.6	47.7	52.9
GNP	506.0	523.3	563.8	594.7	635.7	688.1	753.0	796.3	868.5	935.5
السنة	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
M	58.5	64.0	75.9	94.4	131.9	126.9	155.4	185.8	217.5	260.9
GNP	982.4	1,063.4	1,171.1	1,306.6	1,412.9	1,528.8	1,702.2	1,899.5	2,127.6	2,368.5

المصدر : التقرير الاقتصادى الرئيس ، مكتب حكومة الولايات المتحدة للطباعة ، واشنطن ، ١٩٨٠ صفحة ٢٠٣ .

(أ)

$$\hat{M}_t = -56.13 + 0.13\,GNP \qquad R^2 = 0.98$$
$$(-10.32) \qquad (28.92) \qquad d = 0.65$$

وحيث أن $d = 0,65 < d_L = 1.20$ عند مستوى معنوية 5% مع $n = 20$ و $k' = 1$ (من ملحق ٨)، فهناك دليل على وجود ارتباط ذاتى موجب من الدرجة الأولى .

(ب) لتصحيح الارتباط الذاتى ، يتم إجراء الانحدار الآتى :

$$\hat{M}_t = -20.89 + \underset{(3.58)}{0.72 M_{t-1}} + \underset{(1.36)}{0.15\,\text{GNP}_t} - \underset{(-0.89)}{0.12\,\text{GNP}_{t-1}} \qquad R^2 = 0.99$$

ثم ، باستخدام $\hat{\rho} = 0.72$ (معامل M_{t-1} فى معاملة الانحدار السابقة) ، نحول المتغيرات الأصلية كما سبق الإشارة فى معادلة (٩ - ٣) . المتغيرات الأصلية (M و GNP) والمتغيرات المحولة (M^* و GNP^*) معطاة فى جدول ٩ - ١١ .

$$M^*_{1960} = 23.2\sqrt{1 - 0.72^2} = 16.100 \quad \text{and} \quad Y^*_{1960} = 506.0\sqrt{1 - 0.72^2} = 351.151.$$

باجراء انحدار M^* على GNP^* ، نحصل على :

$$\hat{M}^*_t = \underset{(-5.73)}{-22.43} + \underset{(15.78)}{0.14\,\text{GNP}^*_t} \qquad R^2 = 0.93 \quad d = 2.57$$

وحيث أنه الآن $d = 2.57 > d_U = 1.47$ عند مستوى معنوية 5% مع 20 = n و 1 = k' (من ملحق ٨)، فليس هناك دليل على وجود ارتباط ذاتى . لاحظ أنه بالرغم من أن GNP^*_t يظل لها معنوية عالية ، فإن قيمة t بالنسبة لها أقل من قيمة t للمتغير GNP_t . بالإضافة ، أصبحت $R^2 = 0.93$ الآن مقابل $R^2 = 0.98$ قبل التصحيح بالنسبة للارتباط الذاتى .

جدول ٩ - ١١ الواردات الأمريكية و GNP فى صورتها الأصلية والمحولة

السنة	M	GNP	M^*	GNP*
1960	23.2	506.0	16.100	351.151
1961	23.1	523.3	6.396	158.980
1962	25.2	563.8	8.568	187.024
1963	26.4	594.7	8.256	188.764
1964	28.4	635.7	9.392	207.516
1965	32.0	688.1	11.552	230.396
1966	37.7	753.0	14.660	257.568
1967	40.6	796.3	13.456	254.140
1968	47.7	868.5	18.468	295.164
1969	52.9	935.5	18.556	310.180
1970	58.5	982.4	20.412	308.840
1971	64.0	1063.4	21.880	356.072
1972	75.9	1171.1	29.820	405.452
1973	94.4	1306.6	39.752	463.408
1974	131.9	1412.9	63.932	472.148
1975	126.9	1528.8	31.932	511.512
1976	155.4	1702.2	64.032	601.464
1977	185.8	1899.5	73.912	673.916
1978	217.5	2127.6	83.724	759.960
1979	260.9	2368.5	104.300	836.628

٩ - ١١ يعطى جدول ٩ - ١٢ إجمالى الاستثمار المحلى الخاص ، GPDI و GNP ، كليهما بالبليون دولار ، والرقم القياسى لأسعار المستهلكين ، CPI للولايات المتحدة من ١٩٦٢ إلى ١٩٧٩ (أ) أجر انحدار GPDI على GNP و CPI واختبر وجود ارتباط ذاتى عند مستوى معنوية 5% . (ب) صحح بالنسبة للارتباط الذاتى إن وجد فى (أ)

(أ)
$$\widehat{\text{GPDI}}_t = 125.44 + \underset{(5.47)}{0.37\,\text{GNP}_t} - \underset{(-3.12)}{2.93\,\text{CPI}_t} \qquad R^2 = 0.98 \quad d = 0.72$$

حيث $d = 0.72 < d_L = 1.05$ عند مستوى ثقة 5% بدرجات حرية 18 = n و 2 = k' (من ملحق ٨) ، فهناك دليل على وجود ارتباط ذاتى .

جدول ٩ – ١٢ GPDI و GNP و CPI (بالبليون دولار) فى الولايات المتحدة ، ١٩٦٣ – ١٩٧٩

السنة	1962	1963	1964	1965	1966	1967	1968	1969	1970
GPDI	85.2	90.2	96.6	112.0	124.5	120.8	131.5	146.2	140.8
GNP	563.8	594.7	635.7	688.1	753.0	796.3	868.5	935.5	982.4
CPI	90.6	91.7	92.9	94.5	97.2	100.0	104.2	109.8	116.3
السنة	1971	1972	1973	1974	1975	1976	1977	1978	1979
GPDI	160.0	188.3	220.0	214.6	190.9	243.0	303.3	351.5	386.2
GNP	1,063.4	1,171.1	1,306.6	1,412.9	1,528.8	1,702.2	1,899.5	2,127.6	2,368.5
CPI	121.3	125.3	133.1	147.7	161.2	170.5	181.5	195.4	217.4

المصدر التقرير الإقتصادى للرئيس ، مكتب حكومة الولايات المتحدة للطباعة واشنطن ، ١٩٨٠ ، صفحة ٢٠٣

(ب) للتصحيح لوجود ارتباط ذاتى ، فإننا نجرى أولا الانحدار التالى :

$$\widehat{GPDI}_t = 47.56 + \underset{(3.12)}{0.70}GPDI_{t-1} + \underset{(3.97)}{0.68}GNP_t - \underset{(-2.83)}{0.60}GNP_{t-1} + \underset{(-1.84)}{3.08}CPI_t - \underset{(1.17)}{2.11}CPI_{t-1}$$

$$R^2 = 0.99$$

ومن ثم ، فباستخدام $\hat{\rho} = 0.70$ (معامل $GPDI_{t-1}$ فى الانحدار السابق)، نحول كل المتغيرات الأصلية كما هو موضح فى معادلة (٩ – ٣) . وتظهر المتغيرات الأصلية والمحولة (تميز الأخيرة بعلامة النجمة) فى جدول ٩ – ١٣ .

$$GPDI^*_{1962} = 85.2\sqrt{1 - 0.70^2} = 60.85$$

$$GNP^*_{1962} = 563.8\sqrt{1 - 0.70^2} = 402.63$$

$$CPI^*_{1962} = 90.6\sqrt{1 - 0.70^2} = 64.70$$

بايجاد انحدار $GPDI^*_t$ على GNP^*_t و CPI^*_t نحصل على

$$\widehat{GPDI}_t = 7.19 + \underset{(5.50)}{0.24}GNP^*_t - \underset{(-1.75)}{0.99}CPI^*_t \qquad R^2 = 0.88 \quad d = 1.54$$

وحيث أنه الآن $d = 1.54 > d_U = 1.53$ عند مستوى معنوية 5% مع $n = 20$ و $k' = 2$ (من ملحق ٨)، فليس هناك دليل على وجود ارتباط ذاتى . لاحظ أنه بالرغم من أن GNP^*_t تظل عالية المعنوية ، إلا أن CPI^*_t لا تصبح معنوية . وكذلك فإن R^2 منخفضة .

جدول ٩ – ١٣ GPDI ، GNP و CPI فى الصورة الأصلية والمحولة

السنة	GPDI	GNP	CPI	GPDI*	GNP*	CPI*
1962	85.2	563.8	90.6	60.85	402.63	64.70
1963	90.2	594.7	91.7	30.56	200.04	28.28
1964	96.6	635.7	92.9	33.46	219.41	28.71
1965	112.0	688.1	94.5	44.38	243.11	29.47
1966	124.5	753.0	97.2	46.10	271.33	31.05
1967	120.8	796.3	100.0	33.65	269.70	31.96
1968	131.5	868.5	104.2	46.94	311.09	34.20
1969	146.2	935.5	109.8	54.15	327.55	36.86
1970	140.8	982.4	116.3	38.46	327.55	39.44
1971	160.0	1063.4	121.3	61.44	375.72	39.89
1972	188.3	1171.1	125.3	76.30	426.72	40.39
1973	220.0	1306.6	133.1	88.19	486.83	45.39
1974	214.6	1412.9	147.7	60.60	498.28	54.53
1975	190.9	1528.8	161.2	40.68	539.77	57.81
1976	243.0	1702.2	170.5	109.37	632.04	57.66
1977	303.3	1899.5	181.5	133.20	707.96	62.15
1978	351.5	2127.6	195.4	139.19	797.95	68.35
1979	386.2	2368.5	217.4	140.15	879.18	80.62

٩ – ١٢ يعطى جدول ٩ – ١٤ الإنفاق الاستهلاكى الشخصى ، C ، والدخل الشخصى المتاح Y ، كليهما بالبليون دولار ، للولايات المتحدة من ١٩٦٢ إلى ١٩٧٩ . (أ) أجر انحدار C_t على Y_t واختبر وجود ارتباط داتى . (ب) أجر تصحيحاً بسبب الارتباط الذاتى إن وجد فى (أ) .

جدول ٩ – ١٤ الإنفاق الاستهلاكى والدخل المتاح (بالبليون دولار) : الولايات المتحدة ، ١٩٦٢ – ١٩٧٩ :

السنة	1962	1963	1964	1965	1966	1967	1968	1969	1970
PCI	355.2	374.6	400.4	430.2	464.8	490.4	535.9	579.7	618.8
DPI	383 9	402.8	437.0	472.2	510.4	544.5	588.1	630.4	685.9
السنة	1971	1972	1973	1974	1975	1976	1977	1978	1979
PCI	668.2	733.0	809.9	889.6	979.1	1,089.9	1,210.0	1,350.8	1,509.8
DPI	742.8	801.3	901.7	984.6	1,086.7	1,184.5	1,305.1	1,458.4	1,623.2

المصدر : **التقرير الاقتصادى للرئيس** ، مكتب حكومة الولايات المتحدة للمطبوعات ، واشنطن ، ١٩٨٠ ، صفحة ٢٢٩ .

(أ)

$$\hat{C}_t = -11.40 + 0.93\,Y_t \qquad R^2 = 0.999$$
$$(-1.97) \qquad (144.33) \qquad d = 0.75$$

حيث أن $d = 0.75$ ، فهناك دليل على وجود الارتباط الذاتى عند مستويات المعنوية 5 و 1%

(ب) لعمل تصحيح بسبب الارتباط الذاتى ، نجرى أولا الانحدار الآتى :

$$\hat{C}_t = -25.45 + 1.06\,C_{t-1} + 0.21\,Y_t - 0.16\,Y_{t-1} \qquad R^2 = 0.999$$
$$(5.82) \qquad (1.15) \qquad (-0.89)$$

حيث أن $\hat{\rho} \cong 1$ (معامل C_{t-1} فى المعادلة السابقة) ، فإننا نعيد تقدير الانحدار باستخدام الفروق الأولى للمتغيرات الأصلية (أى انحدار ΔC_t على ΔY_t) مع حذف المقطع ونحصل على

$$\Delta\hat{C}_t = 0.94\,\Delta Y_t \qquad R^2 = 0.96$$
$$(37.79) \qquad d = 2.44$$

قيمة d الجديدة لا تشير إلى أى دليل على وجود ارتباط ذاتى عند مستويات المعنوية 1 أو 5% .

أخطاء فى المتغيرات :

٩ – ١٣ (أ) ماذا يقصد بأخطاء فى المتغيرات ؟ (ب) ما هى المشاكل التى تخلقها الأخطاء فى المتغيرات ؟ (ج) هل هناك اختبارات لاكتشاف وجود أخطاء فى المتغيرات ؟ (د) كيف يمكن تصحيح المشاكل التى يسببها وجود أخطاء فى المغيرات ؟

(أ) تشير أخطاء فى المتغيرات إلى الحالة التى تتضمن فيها متغيرات النموذج أخطاء فى القياس . ومن الممكن أن يكون هذا شائعاً جداً على ضوء الطريقة التى تجمع وتعد بها معظم البيانات .

(ب) أخطاء القياس فى المتغير التابع تدخل فى حد التشويش تاركة تقديرات معالم OLS غير متحيزة (بالرغم من عدم كفاءتها من حيث أن تباينها أكبر من أصغر تباين) . ولكن ، عندما تكون أخطاء القياس فى المتغيرات المفسرة ، فإن هذا يسبب خرق فرض OLS الخامس الخاص باستقلال المتغيرات المستقلة أو المفسرة عن حد الخطأ (أنظر المسألة ٦ – ٤) مما يؤدى إلى تقديرات معالم OLS متحيزة وغير متسقة . فى حالة الانحدار البسيط ، تكون $\hat{b}_1$ متحيزة إلى أدنى ، بينما $\hat{b}_0$ متحيزة إلى أعلى .

(ج) ليس هناك اختبار رسمى للكشف عن وجود أخطاء فى المتغيرات . ولكن يمكن أحياناً أن تعطى النظرية الاقتصادية أو المعرفة بالطريقة التى جمعت بها البيانات إشارة إلى مدى خطورة المشكلة .

(د) واحدة من طرق الحصول على تقديرات معالم OLS متسقة (ولكنها تظل متحيزة وغير كفوء) هى إحلال المتغير المفسر المتضمن أخطاء فى القياس بمتغير آخر له ارتباط عال بهذا المتغير ولكنه مستقل عن حد الخطأ . وفى الواقع العملى ، قد يكون من الصعب العثور على متغير وسيط كهذا ، ولن يكون الإنسان متأكداً أنه سوف يكون مستقلا عن حد الخطأ . والمتغير الوسيط الأكثر شيوعاً هو استخدام القيمة المبطأة للمتغير المفسر محل التساؤل . كما يمكن تصحيح أخطاء القياس فى المتغير المفسر فقط باستخدام المربعات الصغرى المعكوسة . ويتضمن هذا إيجاد انحدار X على Y . فتكون $\hat{b}_0 = -\hat{b}'_0/\hat{b}'_1$ و $\hat{b}_1 = 1/b'_1$ حيث $\hat{b}_0$ ، $\hat{b}_1$ هى تقديرات متسقة للمقطع ومعلمة الميل لانحدار Y_t على X_t .

٩ - ١٤ يعطى جدول ٩ - ١٥ المخزون I ، والمبيعات الفعلية ، S ، وقيمة مفترضة المتغير S تشمل أخطاء فى القياس S' ، كلها بالبليون دولار ، للصناعة التحويلية الأمريكية من ١٩٦٣ إلى ١٩٧٨ . من المفترض أن I و S خالية من أخطاء القياس . (أ) أجر انحدار I_t على S_t (ب) أجر انحدار I_t على S'_t (بافتراض أن S غير متاحة) . ما نوع التحيز الذى ينتج فى التقديرات باستخدام S' بدلا من S ؟ (جـ) استخدم متغيرات وسيطة للحصول على تقديرات معالم متسقة ، على فرض أن S_t ترتبط مع u_t . كيف تقارن تقديرات المعالم هذه مع تلك السابق الحصول عليها فى (ب) ؟

جدول ٩ - ١٥ المخزون والمبيعات (بالبليون دولار) فى الصناعة التحويلية الأمريكية ، ١٩٦٣ - ١٩٧٨ .

Year	1963	1964	1965	1966	1967	1968	1969	1970
I	60.0	63.4	68.2	78.0	84.7	90.6	98.2	101.7
S	35.1	37.3	41.0	44.9	46.5	50.3	53.5	52.8
S'	38.7	41.3	45.6	50.1	51.9	56.3	60.1	59.2
Year	1971	1972	1973	1974	1975	1976	1977	1978
I	102.7	108.3	124.7	157.9	158.2	170.2	180.0	198.0
S	55.9	63.0	73.0	84.8	86.6	98.8	110.8	124.7
S'	62.8	71.0	82.7	96.4	98.5	112.6	126.5	142.7

المصدر : التقرير الاقتصادى للرئيس ، مكتب حكومة الولايات المتحدة للطباعة واشنطن ، ١٩٨٠ ، صفحة ٢٥٦ .

(أ)
$$\hat{I}_t = \underset{(2.10)}{9.65} + \underset{(24.80)}{1.60\,S_t} \qquad R^2 = 0.978 \quad d = 0.74$$

لاحظ أن قيمة d المنخفضة تشير إلى وجود ارتباط ذاتى . وحيث أن الارتباط الذاتى لا ينتج عنه تقديرات متحيزة للمعالم وما يهمنا هنا هو الخطأ فى المتغيرات ، فإننا نمضى بدون تصحيح بسبب الارتباط الذاتى .

(ب)
$$\hat{I}_t = \underset{(2.77)}{12.43} + \underset{(24.79)}{1.38\,S'_t} \qquad R^2 = 0.978 \quad d = 0.74$$

فى وجود أخطاء فى قياس قيمة المبيعات $\hat{b}'_0 > \hat{b}_0$ بينما $\hat{b}'_1 < \hat{b}_1$.

(جـ) باستخدام S'_{t-1} كمتغير وسيط بدلا من S'_t (إذا كان من المعتقد أن S'_t ترتبط مع u_t) ، نحصل على

$$\hat{I}_t = \underset{(1.59)}{8.28} + \underset{(22.66)}{1.58\,S'_{t-1}} \qquad R^2 = 0.975 \quad d = 1.49 \quad r_{S'_t S'_{t-1}} = 0.993$$

لاحظ أن تقديرات المعالم الجديدة أقرب إلى التقديرات الحقيقية عن تلك المقدرة فى (ب) . إن تقديرات المعالم الجديدة هذه ما تزال متحيزة ، ولكنها الآن متسقة (أى أنها تؤول إلى القيم الحقيقية للمعالم مع كبر حجم العينة) . يجب أيضاً ملاحظة أن قيمة d الجديدة تشير إلى عدم وجود ارتباط ذاتى . وطبعاً ، فى الواقع العملى ليس معتاداً أن نعرف أى أخطاء فى القياس قد تكون موجودة (وإلا كان من الممكن تصحيحها قبل إجراء الانحدار) . وأيضاً من الصعب أو من غير الممكن تحديد ما إذا كانت S'_t مرتبطة مع u_t .

٩ – ١٥ باستخدام بيانات جدول ٩ – ١٥ ، (أ) أجر انحدار S_t' على I_t للتغلب على الأخطاء فى قياس S_t . (ب) كيف تقارن هذه النتائج مع تلك فى المسألة ٩ – ١٤ ؟ (جـ) ؟

(أ) حيث أن S_t فقط (أى المتغير المفسر) تتعرض لأخطاء القياس ، فإن المربعات الصغرى المعكوسة هى طريقة أخرى للحصول على تقديرات معالم متسقة . باجراء انحدار S_t' على I_t ، نحصل على

$$\hat{S}_t' = -7.17 + 0.71 I_t \qquad R^2 = 0.978$$
$$(-2.03) \quad (24.79) \qquad d = 0.74$$

$$\hat{b}_0 = -\frac{\hat{b}_0'}{\hat{b}_1'} = -\frac{(-7.17)}{0.71} = 10.10 \quad \text{و} \quad \hat{b}_1 = \frac{1}{\hat{b}_1'} = \frac{1}{0.71} = 1.41$$

حيث b_0 و b_1 تقديرات متسقة (ولكنها لا تزال متحيزة) لمعالم المقطع والميل لانحدار I_t على S_t .

(ب) أن استخدام المربعات الصغرى المعكوسة لا يعطى نتائج فى نفس جودة التقديرات التى أعطتها طريقة المتغير الوسيط (انظر المسألة ٩ – ١٤ (جـ)) . فى حالة المتغير الوسيط . كان تقدير معامل الميل أقرب للقيمة الحقيقية كما تم التخلص من الارتباط الذاتى . ولكن ، النتائج قد تختلف فى حالات أخرى وفى جميع الأحوال ، فإننا فى الواقع العملى كثيراً ما لا نعلم أى نوع من الأخطاء موجود ، وأى نوع من التعديل يناسبها ، وإلى أى حد تقترب المعالم المعدلة من قيم المعالم الحقيقية .

مسائل اضافية

تعدد العلاقات الخطية :

٩ – ١٦ لماذا لا يمكن تقدير دالة الاستهلاك الآتية

$$C_t = b_0 + b_1 Y_{dt} + b_2 Y_{dt-1} + b_3 \Delta Y_{dt} + u_t$$

حيث $\Delta Y_{dt} = Y_{dt} - Y_{dt-1}$ ؟

الإجابة : لأن هناك تعدد خطى تام بين ΔY_{dt} من ناحية و Y_{dt} و Y_{dt-1} من ناحية أخرى . والنتيجة أن هناك ثلاث معادلات طبيعية مستقلة فقط . وأربع معاملات يجب تقديرها وبالتالى لا يكون هناك حل وحيد ممكن .

٩ – ١٧ يعطى جدول ٩ – ١٦ بيانات افتراضية عن الإنفاق الإستهلاكى C ، الدخل المتاح، Y_d والثروة W ، كلها بالألف دولار ، لعينة من 15 أسرة . (أ) أجر انحدار C على Y_d و W ثم أوجد $\bar{R}^2$ و r_{Y_dW} (ب) أجر انحدار C على Y_d فقط (جـ) أجر انحدار C على W فقط (د) ما هى النتائج التى تصل إليها مما سبق فيما يتعلق بالتعدد الخطى ؟

جدول ٩ – ١٦ الإنفاق الاستهلاكى ، الدخل المتاح ، والثروة لعدد 15 أسرة

Family	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C	32	11	15	17	16	13	18	20	14	17	41	17	33	20	18
Y_d	36	12	16	18	17	14	20	23	15	18	50	19	37	22	19
W	144	47	63	70	67	52	79	90	58	70	204	76	149	86	76

الإجابة :

(أ)
$$\hat{C} = 1.54 + 1.41 Y_d - 0.15 W \qquad R^2 = 0.994,\ \bar{R}^2 = 0.993,\ r_{Y_dW} = 0.995$$
$$(1.94) \quad (-0.83)$$

(ب)
$$\hat{C} = 2.13 + 0.80 Y_d \qquad R^2 = 0.994$$
$$(4.98) \quad (46.25)$$

(جـ)
$$\hat{C} = 2.92 + 0.19 W \qquad R^2 = 0.992$$
$$(6.37) \quad (41.46)$$

(د) يوجد تعدد خطى مرتفع .

٩ – ١٨ كيف يمكن استخدام معلومات مسبقة بأن $b_2 = 0.25b_1$ للتغلب على مشكلة التعدد الخطى فى المسألة ٩ – ١٧؟ (ب) أعد تقدير الانحدار فى المسألة ٩ – ١٧ باستخدام المعلومات المسبقة المشار إليها فى (أ) للتغلب على مشكلة التعدد الخطى .
(جـ) ما هى قيمة $\hat{b}_1$ ؟ $\hat{b}_2$ ؟

الإجابة : (أ) بتقدير $C = b_0 + b_1 Z$ ، حيث $Z = Y_d + 0.25W$

(ب)
$$\hat{C} = \underset{(5.75)}{2.53} + \underset{(44.10)}{0.39Z} \qquad R^2 = 0.993$$

(جـ)
$$\hat{b}_1 = 0.39 \text{ and } \hat{b}_2 = 0.10$$

٩ – ١٩ يعطى جدول ٩ – ١٧ إجمالى التكوين الرأسمالى الثابت Y_i ، والمبيعات X_i كليهما بالألف دولار لعدد 35 شركة فى إحدى الصناعات . أجر انحدار Y_i على X_i (أ) لكل البيانات (ب) لعدد 14 مشاهدة الأولى فقط وسجل مجموع مربعات الخطأ ، ESS_1 ، (جـ) لعدد 14 مشاهدة الأخيرة فقط ، وسجل مجموع مربعات الخطأ ، ESS_2 .
(د) اختبر وجود اختلاف التباين .

جدول ٩ – ١٧ إجمالى التكوين الرأسمالى الثابت والمبيعات لعدد 35 شركة

المبيعات	إجمالى التكوين الرأسمالى الثابت						
100	31.2	31.2	30.9	30.7	30.5	30.5	30.2
150	33.4	33.4	32.8	32.3	31.9	31.5	31.5
200	37.8	37.4	37.4	36.9	36.3	35.7	35.1
250	43.2	42.9	42.1	40.8	40.2	39.1	38.4
300	48.5	47.7	46.5	45.9	45.2	44.9	44.3

الإجابة : (أ)
$$\hat{Y}_i = \underset{(28.50)}{21,637} + \underset{(22.00)}{0.079X_i} \qquad R^2 = 0.94$$

(ب)
$$\hat{Y}_i = \underset{(31.51)}{27,429} + \underset{(4.85)}{0.033X_i} \qquad R^2 = 0.66 \quad ESS_1 = 4.897$$

(جـ)
$$\hat{Y}_i = \underset{(2.99)}{15,029} + \underset{(5.71)}{0.104X_i} \qquad R^2 = 0.73 \quad ESS_2 = 34.694$$

(د) حيث أن $ESS_2/ESS_1 = 7.08$ تتجاوز $F_{11,11} = 2.82$ عند مستوى معنوية 5% ، فإن هناك اختلافاً فى التباين .

٩ – ٢٠ بافتراض أن تباين الخطأ يتناسب مع X_i^2 فى المسألة ٩ – ١٩ : (أ) صحح لاختلاف التباين ، (ب) ما هى القيمة الجديدة للمقطع ، وقيمة معلمة الميل الجديدة المرتبطة بالمتغير X_i ؟ كيف تقارن بالقيم المناظرة قبل التحويل ؟

الإجابة : (أ)
$$\hat{Y}_i/X_i = \underset{(20.41)}{0.074} + \underset{(42.16)}{23,187\,(1/X_i)} \qquad R^2 = 0.98$$

(ب) القيمة الجديدة للمقطع هى 23, 187 (بدلا من 21, 637) ومعلمة الميل الجديدة الخاصة بالمتغير X_i هى الآن 0.074 (بدلا من 0.079) .

٩ – ٢١ يعطى جدول ٩ – ١٨ مستوى إجمالى التكوين الرأسمالى الثابت Y ، والمبيعات X_1 ، كليهما بالألف دولار ، ورقم قياسى للانتاجية X_2 ، لعدد 25 شركة فى إحدى الصناعات . من المتوقع أن Y سوف ترتبط مباشرة مع كل من X_1 و X_2 .

أجر انحدار Y على X_1 و X_2 (أ) لكل العينة ، (ب) لعدد 14 مشاهدة المقابلة لأصغر قيم X_2 وسجل ESS_1 و (جـ) لعدد 14 مشاهدة المقابلة لأكبر قيم X_2 وسجل ESS_2 . (د) اختبر وجود اختلاف التباين .

الإجابة : (أ)

$$\hat{Y} = 12{,}089 + \underset{(2.53)}{0.017X_1} + \underset{(8.93)}{1.608X_2} \qquad R^2 = 0.99$$

(ب)

$$\hat{Y} = 33{,}332 + \underset{(3.91)}{0.044X_1} - \underset{(-0.99)}{0.784X_2} \qquad R^2 = 0.95 \qquad ESS_1 = 0.658$$

(جـ)

$$\hat{Y} = 5{,}874 + \underset{(0.30)}{0.010X_1} + \underset{(2.56)}{2.115X_2} \qquad R^2 = 0.99 \qquad ESS_2 = 2.126$$

جدول ٩ – ١٨ إجمالى التكوين الرأسمالى الثابت ، المبيعات ، والإنتاجية فى 35 شركة

الشركة	1	2	3	4	5	6	7	8	9	10	11	12
Y	30.9	31.5	43.2	36.9	44.3	30.5	32.3	42.9	31.2	39.1	35.7	40.8
X_1	135	150	300	225	310	105	170	285	145	250	205	275
X_2	10.3	10.8	16.4	12.9	16.7	10.0	10.9	15.9	10.6	14.6	12.1	15.5
الشركة	13	14	15	16	17	18	19	20	21	22	23	24
Y	31.2	42.1	32.8	36.3	37.4	30.5	33.4	37.4	44.9	33.4	45.2	30.2
X_1	140	280	180	215	235	110	190	230	315	195	320	100
X_2	10.5	15.6	10.9	12.5	13.8	10.0	11.1	13.1	17.1	11.3	17.3	9.9
الشركة	25	26	27	28	29	30	31	32	33	34	35	
Y	45.9	46.8	35.1	40.2	47.9	30.7	38.1	49.3	31.9	37.8	31.5	
X_1	330	345	200	260	350	120	250	355	165	245	150	
X_2	17.5	17.9	11.5	14.9	18.3	10.1	14.1	18.5	10.8	13.9	10.7	

(د) حيث أن $ESS_2/ESS_1 = 3.23$ تجاوز $F_{11,11} = 2.82$ عند مستوى معنوية 5% ، فهناك اختلاف فى التباين .

٩ – ٢٢ بافتراض أن تباين الخطأ يتناسب مع X_2^2 فى المسألة ٩ – ٢١ (أ) صحح لاختلاف التباين . (ب) ما هى القيمة الجديدة للمقطع ومعاملات الميل للمتغيرات X_1 و X_2 ؟ كيف تقارن بالقيم [illegible] قبل التحويل ؟

الإجابة : (أ)

$$\hat{Y}/X_2 = \underset{(10.53)}{1.622} + \underset{(2.85)}{0.016(X_1/X_2)} + 12{,}200(1/X_2) \qquad R^2 = 0.94$$

(ب) قيمة المقطع الجديدة هى 12,000 (بدلا من 12.089) ، بينما القيمة الجديدة لمعامل ميل X_1 هى 0.016 (بدلا من 0.017) ولمعامل ميل X_2 هى 1.622 (بدلا من 1.608) .

الارتباط الذاتى :

٩ – ٢٣ يعطى جدول ٩ – ١٩ إنفاق قطاع الأعمال على المصانع والأجهزة الجديدة للمرافق العامة Y ، ومستوى GNP و X_1 ، كليهما بالبليون دولار ، والرقم القياسى لأسعار السلع X_2 للولايات المتحدة من ١٩٦٢ إلى ١٩٧٩ . (أ) أجر انحدار Y على X_1. هل هناك دليل على وجود ارتباط ذاتى عند مستويات معنوية 5 و 1% ؟ (ب) أجر انحدار Y_t على Y_{t-1} ، X_{1t} و X_{1t-1} . ما هى قيمة ρ ؟ (جـ) أجر انحدار Y_t^* على X_{1t}^* للتصحيح بسبب الارتباط الذاتى ، حيث Y_t^* و X_{1t}^* هى المتغيرات المحولة . هل هناك دليل على وجود ارتباط ذاتى عند مستوى معنوية 1% ؟ عند مستوى معنوية 5% ؟

جدول ٩ – ١٩ انفاق قطاع الأعمال على المصانع والأجهزة الجديدة للمرافق العامة ، GNP (بالبليون دولار) ، والرقم القياسى لأسعار السلع : الولايات المتحدة ، ١٩٦٢ – ١٩٧٩ .

السنة	1962	1963	1964	1965	1966	1967	1968	1969	1970
Y	4.9	5.0	5.5	6.3	7.4	8.7	10.2	11.6	13.1
X_1	563.8	594.7	635.7	688.1	753.0	796.3	868.5	935.5	982.4
X_2	92.8	93.6	94.6	95.7	98.2	100.0	103.7	108.4	113.5
السنة	1971	1972	1973	1974	1975	1976	1977	1978	1979
Y	15.3	17.0	18.7	20.6	20.1	22.3	25.8	29.5	33.2
X_1	1,063.4	1,171.1	1,306.6	1,412.9	1,528.8	1,702.2	1,899.5	2,127.6	2,368.5
X_2	117.4	120.9	129.9	145.5	158.4	165.2	174.7	187.1	208.4

المصدر : التقرير الاقتصادى للرئيس ، مكتب حكومة الولايات المتحدة للطباعة واشنطن ، ١٩٨٠ صفحة ٢٥٥ ،صفحة ٢٦٢ .

الإجابة : (أ)
$$\hat{Y}_t = -3.462 + 0.016X_{1t} \qquad R^2 = 0.98$$
$$(-5.09) \qquad (30.23) \qquad d = 0.38$$

حيث أن $d = 0.38$ ، فهناك دليل على وجود ارتباط ذاتى عند كل من مستويات المعنوية 5 و 1% .

(ب)
$$\hat{Y}_t = -0.342 + 0.821Y_{t-1} + 0.016X_{1t} - 0.013X_{1t-1} \qquad R^2 = 0.99$$
$$(4.77) \qquad (1.33) \qquad (-0.89) \qquad \hat{\rho} \cong 0.82$$

(جـ)
$$\hat{Y}_t^* = -0.446 + 0.015X_{1t}^* \qquad R^2 = 0.92$$
$$(-1.01) \qquad (13.64) \qquad d = 1.36$$

لا يوجد دليل على وجود ارتباط ذاتى عند مستوى المعنوية 1% ، ولكن الاختبار غير حاسم عند مستوى المعنوية 5%.

٩ – ٢٤ باستخدام بيانات جدول ٩ – ١٩ (أ) أجر انحدار Y_t على X_{1t} و X_{2t} . هل هناك دليل على وجود ارتباط ذاتى عند مستوى معنوية 5 و 1% ؟ (ب) إذا وجد دليل فى (أ) على وجود ارتباط ذاتى ، اوجد قيمة ρ التى يجب استخدامها لتحويل المتغيرات لتصحيح الارتباط الذاتى (جـ) إذا وجد دليل على وجود ارتباط ذاتى فى (أ) ، أجر انحدار Y_t^* على X_{1t}^* و X_{2t}^* لتصحيح الارتباط الذاتى . هل هناك دليل على ارتباط باق عند مستوى معنوية 1% ؟ عند مستوى معنوية 5% ؟

الإجابة : (أ)
$$\hat{Y}_t = 4.113 + 0.026X_{1t} - 0.152X_{2t} \qquad R^2 = 0.99$$
$$(6.12) \qquad (-2.39) \qquad d = 0.62$$

هناك دليل على وجود ارتباط ذاتى عند مستوى معنوية 5 و 1% (ب) $\rho = 0.62$ (جـ)
$$Y_t^* = 0.196 + 0.020X_{1t}^* - 0.073X_{2t}^* \qquad R^2 = 0.97$$
$$(10.95) \qquad (-2.82) \qquad d = 1.33$$

ليس هناك دليل على وجود ارتباط ذاتى عند مستوى معنوية 1% ، ولكن الاختبار غير حاسم عند مستوى معنوية 5% .

٩ – ٢٥ باستخدام بيانات جدول ٩ – ١٩ ، (أ) أجر انحدار ΔY_t على ΔX_{1t} ، (ب) هل هناك دليل على وجود ارتباط ذاتى عند مستوى معنوية 1 و 5% ؟ (جـ) ماذا يبرر هذا التحويل للمتغيرات ؟ هل يظل هذا التحويل له ما يبرره إذا دخلت كل من X_{1t} و X_{2t} فى الانحدار ؟

الإجابة : (أ)
$$\Delta Y_t = 0.015\Delta X_{1t} \qquad R^2 = 0.63$$
$$(10.86) \qquad d = 1.55$$

(ب) ليس هناك دليل الآن على وجود ارتباط ذاتى عند مستويات معنوية 1% و 5% . ولكن هناك خطأ فى تحديد النموذج لأن X_{2t} غير داخلة فى الانحدار . (جـ) هذا التحويل له ما يبرره فقط عندما يكون X_{1t} وحدها فى الانحدار ، لأن فى هذه الحالة قيمة ρ قريبة من 1 $(\hat{\rho} = 0.82)$ (انظر المسألة ٩ – ٢٣ (ب)) . وهذا التحويل مبرراته أقل عندما تدخل X_{1t} و X_{2t} فى الانحدار لأنه عندئذ تكون $\hat{\rho} = 0.62$ فقط (انظر المسألة ٩ – ٢٤ (ب)) .

أخطاء فى المتغيرات :

٩ – ٢٦ يعطى جدول ٩ – ٢٠ المخزون Y ، الشحنات الفعلية ، X ، وقيم افتراضية للمتغير X تتضمن أخطاء فى القياس X'، كلها بالبليون دولار ، فى صناعة السلع المعمرة الأمريكية من ١٩٦٣ إلى ١٩٧٨ . من المفترض أن Y و X خالية من أخطاء القياس . (أ) أجر انحدار Y_t على X_t (ب) أجر انحدار Y_t على X_t' (بافتراض أن X غير متاحة) . ما نوع التحيز الناتج فى التقديرات باستخدام X' بدلا من X ؟ (جـ) استخدم متغيراً وسيطاً للحصول على تقديرات معالم متسقة ، على فرض أن X_t مرتبطة مع u_t . كيف تقارن تقديرات المعالم هذه مع تلك فى (ب) .

الإجابة : (أ)

$$\hat{Y}_t = \underset{(0.78)}{3.01} + \underset{(19.88)}{2.03X_t} \qquad R^2 = 0.966 \quad d = 0.78$$

(ب)

$$\hat{Y}_t = \underset{(1.78)}{6.54} + \underset{(19.90)}{1.75X_t'} \qquad R^2 = 0.966 \quad d = 0.79$$

جدول ٩ – ٢٠ المخزون والشحنات (بالبليون دولار) فى صناعة السلع المعمرة الأمريكية ، ١٩٦٣ – ١٩٧٨ .

السنة	1963	1964	1965	1966	1967	1968	1969	1970
Y	35.9	38.5	42.3	49.9	55.0	58.9	64.7	66.8
X	18.3	19.6	22.2	24.6	25.3	27.7	29.5	28.2
X'	19.2	20.7	23.8	26.5	27.3	30.1	32.2	30.7
السنة	1971	1972	1973	1974	1975	1976	1977	1978
Y	66.3	70.3	81.4	101.9	101.8	109.1	115.6	129.2
X	30.0	34.0	39.7	44.3	43.7	50.7	58.0	66.5
X'	32.8	37.4	44.1	49.4	48.7	56.8	65.3	75.1

المصدر : التقرير الاقتصادى للرئيس ، مكتب حكومة الولايات المتحدة للطباعة ، واشنطن ، ١٩٨٠ ، صفحة ٢٥٧ .

فى وجود خطأ فى قياس قيمة الشحنات ، $b_0' > b_0$، بينما $b_1' < b_1$. (جـ) باستخدام X_{t-1}' كمتغير وسيط للمتغير X_t' ، نحصل على :

$$\hat{Y}_t = \underset{(0.81)}{2.55} + \underset{(25.11)}{2.04X_{t-1}'} \qquad R^2 = 0.980 \quad d = 1.68 \quad r_{X_t'X_{t-1}'} = 0.987$$

تقديرات المعالم الجديدة أقرب للمعالم الحقيقية من تلك السابق الحصول عليها فى (ب) .

٩ - ٢٧ باستخدام بيانات جدول ٩ - ٢٠ ، (أ) أجر انحدار X_t' على Y_t للتغلب على أخطاء القياس فى X_t . متى تكون هذه الطريقة مناسبة ؟ (ب) كيف تقارن النتائج فى (أ) بتلك فى مسألة ٩ - ٢٦ (جـ) ؟

الإجابة : (أ)

$$\hat{X}_t' = -2.29 + 0.55\,Y_t \qquad R^2 = 0.966$$
$$(-1.03)\quad(19.90) \qquad d = 0.80$$

التقديرات المتسقة لمعالم انحدار Y_t على X_t' هى $b_0 = 4.16$ و $b_1 = 1.82$ وتكون المربعات الصغرى المعكوسة مناسبة عندما يتعرض المتغير المفسر فقط لأخطاء القياس .

(ب) استخدام المربعات الصغرى المعكوسة لا يعطى نتائج بنفس جودة النتائج الناجمة عن استخدام طريقة المتغير الوسيط (انظر المسألة ٩ - ٢٦ (جـ)) .

الفصل العاشر

طرق المعادلات الآنية

١٠-١ نماذج المعادلات الآنية

عندما يكون المتغير التابع فى معادلة ما متغيراً مفسراً فى معادلة أخرى ، يكون لدينا **نظام أو نموذج معادلات آنية** . المتغيرات التابعة فى نظام معادلات آنية تسمى أيضاً **بالمتغيرات الداخلية** . بينما تسمى المتغيرات التى تحددها عوامل خارج النموذج **بالمتغيرات الخارجية** . وهناك معادلة **سلوكية أو هيكلية** لكل متغير داخلى فى النظام (أنظر مثال ١) . واستخدام OLS لتقدير المعادلات الهيكلية يؤدى إلى تقديرات معالم متحيزة وغير متسقة . ويشار إلى هذا **بتحيز المعادلات الآنية** . وللحصول على تقديرات معالم متسقة ، يجب الحصول أولا على **معادلات الشكل المختزل** للنموذج . وهذه المعادلات تعبر عن كل متغير داخلى فى النظام كدالة فقط فى المتغير الخارجى للنموذج (أنظر مثال ٢) .

مثال ١ – المعادلتان الآتيتان تمثلان نموذجاً كلياً بسيطاً

$$M_t = a_0 + a_1 Y_t + u_{1t}$$
$$Y_t = b_0 + b_1 M_t + b_2 I_t + u_{2t}$$

حيث M_t هى عرض النقود فى الفترة t ، Y هى الدخل ، و I هى الاستثمار . وحيث أن M تعتمد على Y فى المعادلة الأولى وتعتمد Y على M (وكذلك I) فى المعادلة الثانية ، M و Y تتحدان معاً ، فإن لدينا نموذج معادلات آنية . إن M و Y متغيران داخليان ، بينما I متغير خارجى أى يتحدد خارج النموذج . والتغير فى u_{1t} يؤثر على M_t فى المعادلة الأولى . وهذا بدوره يؤثر على Y_t فى المعادلة الثانية . وكنتيجة يكون Y و u_{1t} مترابطين ، مؤدياً إلى تقديرات OLS متحيزة وغير متسقة لمعادلة M (و Y).

مثال ٢ – يمكن اشتقاق معادلة الشكل المختزل الأولى بالتعويض بالمعادلة الثانية فى المعادلة الأولى وإعادة الترتيب .

$$M_t = a_0 + a_1(b_0 + b_1 M_t + b_2 I_t + u_{2t}) + u_{1t}$$
$$M_t = \frac{a_0 + a_1 b_0}{1 - a_1 b_1} + \frac{a_1 b_2}{1 - a_1 b_1} I_t + \frac{u_{1t} + a_1 u_{2t}}{1 - a_1 b_1}$$

أو

$$M_t = \pi_0 + \pi_1 I_t + v_{1t}$$

ويمكن اشتقاق معادلة الشكل المختزل الثانية بالتعويض بالمعادلة الأولى فى المعادلة الثانية وإعادة الترتيب :

$$Y_t = b_0 + b_1(a_0 + a_1 Y_t + u_{1t}) + b_2 I_t + u_{2t}$$
$$Y_t = \frac{a_0 b_1 + b_0}{1 - a_1 b_1} + \frac{b_2}{1 - a_1 b_1} I_t + \frac{b_1 u_{1t} + u_{2t}}{1 - a_1 b_1}$$

أو

$$Y_t = \pi_2 + \pi_3 I_t + v_{2t}$$

١٠-٢ التمييز

يشير التمييز إلى إمكانية حساب المعالم الهيكلية لنموذج المعادلات الآنية من معالم الشكل المختزل . وتكون معادلة ما فى نظام مميزة بالضبط إذا كان عدد المتغيرات الخارجية المستبعدة من المعادلة مساوياً لعدد المتغيرات الداخلية فيها ناقصاً 1 .

ولكن ، تكون معادلة ما فى نظام **زائدة التمييز** (**أو ناقصة التمييز**) إذا كان عدد المتغيرات الخارجية المستبعدة من المعادلة أكبر من (أو أصغر من) عدد المتغيرات الداخلية الداخلة فى المعادلة 1 (انظر مثال ٣) . وبالرغم من أن هذا شرط ضرورى وليس كافيا للتمييز ، فانه عادة ما يعطى الإجابة الصحيحة (انظر المسألة ١٠-٥) ويمكن حساب معاملات هيكلية وحيدة من معاملات الشكل المختزل فقط للمعادلة المميزة بالضبط (انظر مثال ٤) .

مثال ٣ — معادلة عرض النقود ، M فى مثال ١ مميزة بالضبط لأنها تستبعد متغيراً خارجاً واحداً ، 1 ، وتتضمن متغيرين داخليين ، M و Y . ولكن معادلة الدخل Y ، ناقصة التمييز لأنها لا تستبعد أى متغيرات خارجية . وإذا تضمنت المعادلة الثانية هذه متغيرا خارجياً خارجيا إضافياً G (الإنفاق الحكومى) ، فإن المعادلة الأولى ، معادلة M ، تكون زائدة التمييز ، لأن عدد المتغيرات الخارجية المستبعدة فى هذه الحالة يزيد عن عدد المتغيرات الداخلية ناقصاً 1 .

مثال ٤ — يمكن حساب قيمة وحيدة للمعالم الهيكلية للمعادلة M المميزة بالضبط فى مثال ١ من معالم الشكل المختزل فى مثال ٢ كالآتى :

$$a_0 = \pi_0 - a_1\pi_2 = \frac{a_0(1 - a_1b_1)}{1 - a_1b_1} \quad و \quad a_1 = \frac{\pi_1}{\pi_3} = \frac{\dfrac{a_1b_2}{1 - a_1b_1}}{\dfrac{b_2}{1 - a_1b_1}}$$

١٠-٣ التقدير : المربعات الصغرى غير المباشرة

المربعات الصغرى غير المباشرة (ILS) هى طريقة لحساب قيمة المعالم الهيكلية للمعادلات المميزة بالضبط . وتتضمن ILS استخدام OLS لتقدير معادلات الشكل المختزل للنظام ثم استخدام المعاملات المقدرة لحساب المعاملات الهيكلية . ولكن ، ليس من السهل حساب الأخطاء المعيارية للمعالم الهيكلية ، كما لا يمكن استخدام ILS فى حالات التمييز الزائد .

مثال ٥ — يعطى جدول ١٠-١ عرض النقود (M = العملة زائداً الودائع تحت الطلب) GNP ، Y ، إجمالى الاستثمار المحلى الخاص ، I ، ومشتريات الحكومة من السلع والخدمات G ، كلها بالبليون دولار ، للولايات المتحدة من ١٩٦٢ إلى ١٩٧٩ (سوف تستخدم G فى مثال ٦) . معادلات الشكل المختزل المقدرة لمثال ٢ هى :

جدول ١٠-١ عرض النقود ، GNP ، الاستثمار والإنفاق الحكومى (ببلايين الدولارات) : الولايات المتحدة ، ١٩٦٢ ، ١٩٧٩ .

السنة	1962	1963	1964	1965	1966	1967	1968	1969	1970
M	150.9	156.5	163.7	171.4	175.8	187.4	202.5	209.0	219.7
Y	563.8	594.7	635.7	688.1	753.0	796.3	868.5	935.5	982.4
I	85.2	90.2	96.6	112.0	124.5	120.8	131.5	146.2	140.8
G	118.0	123.7	129.8	138.4	158.7	180.2	198.7	207.9	218.9
السنة	1971	1972	1973	1974	1975	1976	1977	1978	1979
M	233.9	255.3	270.5	283.2	295.4	313.8	338.7	361.5	382.1
Y	1,063.4	1,171.1	1,306.6	1,412.9	1,528.8	1,702.2	1,899.5	2,127.6	2,368.5
I	160.0	188.3	220.0	214.6	190.9	243.0	303.3	351.5	386.2
G	233.7	253.1	269.5	302.7	338.4	361.3	396.2	435.6	476.1

المصدر : التقرير الاقتصادى للرئيس ، مكتب الولايات المتحدة للطباعة واشنطن ، ١٩٨٠ صفحة ٢٧١ ، صفحة ٢٠٣ .

$$\hat{M}_t = \underset{(9.72)}{95.8602} + \underset{(16.48)}{0.8004 I_t} \qquad R^2 = 0.944$$

$$\hat{Y}_t = \underset{(1.41)}{75.7767} + \underset{(22.92)}{6.0608 I_t} \qquad R^2 = 0.970$$

$$\hat{a}_1 = \frac{\hat{\pi}_1}{\hat{\pi}_3} = \frac{0.8004}{6.0608} = 0.1321$$

$$\hat{a}_0 = \hat{\pi}_0 - a_1 \hat{\pi}_2 = 95.8602 - 0.1321(75.7767) = 85.8501$$

وعليه تكون معادلة M فى مثال ١ المقدرة باستخدام ILS .

$$\hat{M}_t = 85.8501 + 0.1321\, Y_t$$

ونفس المعادلة مقدرة (خطأ) باستخدام OLS هى

$$\hat{M}_t = \underset{(16.62)}{84.7943} + \underset{(33.95)}{0.1330 Y_t} \qquad R^2 = 0.986$$

١٠ـ٤ التقدير : المربعات الصغرى على مرحلتين

المربعات الصغرى على مرحلتين (2SLS) هى طريقة لتقدير معالم هيكلية متسقة للمعادلات زائدة التمييز (بالنسبة للمعادلات المميزة بالضبط ، تعطى 2SLS نفس نتائج ILS ولكنها تعطى أيضاً الأخطاء المعيارية للمعالم الهيكلية المقدرة) . وتتضمن 2SLS إجراء انحدار كل متغير داخلى على كل المتغيرات الخارجية فى النظام ثم تستخدم القيم المتوقعة للمتغيرات الداخلية لتقدير المعادلات الهيكلية للنموذج .

مثال ٦ – إذا تضمنت المعادلة الثانية ، معادلة Y ، فى مثال G (الإنفاق الحكومى) كتغير مفسر إضافى ، تصبح المعادلة الأولى ، معادلة M ، زائدة التمييز (انظر مثال ٣) ويمكن تقديرها باستخدام 2 SLS . المرحلة الأولى هى

$$\hat{Y}_t = \underset{(-1.72)}{-27.1686} + \underset{(6.26)}{1.8481 I_t} + \underset{(14.68)}{3.4748 G_t} \qquad R^2 = 0.998$$

المرحلة الثانية هى :

$$M_t = \underset{(18.74)}{84.3989} + \underset{(38.53)}{0.1333 Y_t} \qquad R^2 = 0.989$$

$\hat{a}_1 = 0.1333$ هى تقدير متسق للمعلمة a_1 .

مسائل محلولة

نماذج المعادلات الآنيـة :

١٠ – ١ ماذا يقصد بالآتى : (أ) نموذج أو نظام المعادلات الآنية ؟ (ب) المتغيرات الداخلية ؟ (جـ) المتغيرات الخارجية (د) المعادلات الهيكلية ؟ (هـ) تميز المعادلات الآنية ؟ (و) معادلات الشكل المختزل ؟

(أ) يشير نموذج أو نظام المعادلات الآنية إلى الحالة التى يكون فيها متغير تابع فى معادلة أو أكثر هو متغير مفسر فى معادلة أخرى فى النظام . أى أن قيم Y لا تتحدد فقط عن طريقة قيم X ، ولكن بعضاً من قيم X تتحدد بدورها عن طريق قيم Y بحيث أن قيم Y وقيم X تتحدد آنياً معاً .

(ب) المتغيرات الداخلية هي المتغيرات التابعة في نظام من المعادلات الآنية . وهذه هي المتغيرات التي يحددها النظام ، بالرغم من أنها تظهر أيضاً كمتغيرات مفسرة في بعض معادلات النظام .

(جـ) المتغيرات الخارجية هي تلك المتغيرات التي تتحدد خارج النموذج ، وتتضمن هذه أيضاً المتغيرات الداخلية المبطأة ، حيث أن قيمها تكون معلومة فعلا في أي فترة زمنية معينة . وأحيانا تسمى المتغيرات الخارجية والمتغيرات الداخلية المبطأة بالمتغيرات المحددة سلفاً .

(د) المعادلات الهيكلية أو السلوكية تصف هيكل اقتصاد ما أو سلوك بعض الوكلاء الاقتصاديين مثل المستهلكين أو المنتجين . وهناك معادلة هيكلية واحدة لكل متغير داخلي في النظام . وتسمى معاملات المعادلات الهيكلية بالمعالم الهيكلية وتعبر عن الأثر المباشر لكل متغير مفسر على المتغير التابع .

(هـ) يشير تحيز المعادلات الآنية إلى التقدير الزائد أو التقدير الناقص للمعالم الهيكلية التي يتم الحصول عليها عند تطبيق OLS على المعادلات الهيكلية لنموذج المعادلات الآنية . وينتج هذا التحيز لأن هذه المتغيرات الداخلية في النظام والتي تظهر أيضاً كمتغيرات مفسرة ترتبط مع حدود الخطأ ، وبالتالي تخرق الفرض الخامس من فروض OLS (انظر المسألة ٦ - ٤) .

(و) معادلات الشكل المختزل يتم الحصول عليها بحل نظام المعادلات الهيكلية بحيث يعبر عن كل متغير داخلي في النظام كدالة فقط في المتغيرات الخارجية أو المحددة سلفاً في النظام . وحيث أن المتغيرات الخارجية للنظام لا ترتبط مع حدود الخطأ فإن OLS تعطى تقديرات متسقة لمعالم الشكل المختزل . وتقيس هذه إجمالي الآثار المباشرة وغير المباشرة للتغير في المتغيرات الخارجية على المتغيرات الداخلية ويمكن استخدامها للحصول على تقديرات معالم هيكلية متسقة .

١٠ - ٢ تمثل المعادلتان الهيكليتان التاليتان نموذج عرض - طلب بسيط :

$$Q_t = a_0 + a_1P_t + a_2Y_t + u_{1t}, \quad a_1 < 0 \quad \text{and} \quad a_2 > 0 \qquad : \text{الطلب}$$
$$Q_t = b_0 + b_1P_t + u_{2t} \qquad b_1 > 0 \qquad : \text{العرض}$$

حيث Q هي الكمية ، P هي السعر ، Y دخل المستهلك . من المفترض أن كل الكمية المعروضة تباع على نهاية العام ومن ثم فإن Q_t تمثل كلا من الكمية المباعة والمشتراة خلال العام t (أ) لماذا يعتبر هذا نموذج معادلات آنية ؟ (ب) ما هي المتغيرات الداخلية والمتغيرات المتغيرات الخارجية في النظام ؟ (جـ) لماذا يؤدى استخدام OLS في تقدير معادلتي العرض والطلب إلى تقديرات معالم متحيزة وغير متسقة ؟

(أ) يمثل نموذج العرض - الطلب المعطى نظام معادلات آنية للسوق لأن Q و P تتحددان معاً وبالتبادل . فإذا كان السعر أقل من سعر التوازن ، فإن الكمية المطلوبة تتجاوز الكمية المعروضة ، والعكس بالعكس . عند التوازن ، يقطع منحنى الطلب (السالب الميل) منحنى العرض (الموجب الميل) ويحددان معاً أو آنياً قيم Q و P (التوازنية) .

(ب) المتغيرات الداخلية للنموذج هي Q و P . هذه هي المتغيرات التي تتحدد داخل النموذج . Y هي المتغير الخارجي الوحيد في النموذج (أي أنه يتحدد خارج النموذج) .

(جـ) حيث أن المتغير الداخلي P هو أيضاً متغير مفسر في كل من معادلتي العرض والطلب ، فإن P ترتبط مع u_{1t} في معادلة الطلب ومع u_{2t} في معادلة العرض . ويخرق هذا فرض OLS الخامس ، الذى يتطلب أن يكون المتغير

المفسر غير مرتبط مع حد الخطأ . وكنتيجة ، فإن تقدير معادلتى العرض والطلب باستخدام OLS يؤدى إلى تقديرات معالم ليس فقط متحيزة ولكن أيضاً غير متسقة (أى أنها لا تؤول إلى المعالم الحقيقية مع زيادة حجم العينة) .

١٠ - ٣ (أ) أوجد معادلات الشكل المختزل المناظرة للمعادلات الهيكلية فى المسألة ١٠ - ٢ . (ب) لماذا تكون معادلات الشكل المختزل هذه مهمة ؟ ماذا تقيس معاملات الشكل المختزل فى نموذج السوق هذا ؟

(أ) لإيجاد معادلات الشكل المختزل ، يتم حل المعادلات الهيكلية فى المسألة ١٠ - ٢ بالنسبة لكل من Q و P (المتغيرات الداخلية) كدالة فى Y (المتغير الخارجى) فقط بتحويل معادلة العرض إلى دالة فى P وبالتعويض فى معادلة الطلب ، نحصل على :

$$P_t = \frac{1}{b_1}(Q_t - b_0 - u_{2t})$$

$$Q_t = a_0 + \frac{a_1}{b_1}(Q_t - b_0 - u_{2t}) + a_2 Y_t + u_{1t}$$

$$Q_t\left(\frac{b_1 - a_1}{b_1}\right) = \left(\frac{a_0 b_1 - a_1 b_0}{b_1}\right) + a_2 Y_t + \left(\frac{b_1 u_{1t} - a_1 u_{2t}}{b_1}\right)$$

$$Q_t = \left(\frac{a_0 b_1 - a_1 b_0}{b_1 - a_1}\right) + \left(\frac{b_1 a_2}{b_1 - a_1}\right) Y_t + \left(\frac{b_1 u_{1t} - a_1 u_{2t}}{b_1 - a_1}\right)$$

$$Q_t = \pi_0 + \pi_1 Y_t + v_{1t}$$

حيث

$$\pi_0 = \frac{a_0 b_1 - a_1 b_0}{b_1 - a_1} \quad \pi_1 = \frac{b_1 a_2}{b_1 - a_1} \quad v_{1t} = \frac{b_1 u_{1t} - a_1 u_{2t}}{b_1 - a_1}$$

بالتعويض بمعادلة الطلب فى معادلة العرض كدالة P ، نحصل على :

$$P_t = \frac{1}{b_1}(a_0 + a_1 P_t + a_2 Y_t + u_{1t} - b_0 - u_{2t})$$

$$P_t\left(\frac{b_1 - a_1}{b_1}\right) = \frac{1}{b_1}(a_0 + a_2 Y_t + u_{1t} - b_0 - u_{2t})$$

$$P_t = \left(\frac{a_0 - b_0}{b_1 - a_1}\right) + \left(\frac{a_2}{b_1 - a_1}\right) Y_t + \left(\frac{u_{1t} - u_{2t}}{b_1 - a_1}\right)$$

$$P_t = \pi_2 + \pi_3 Y_t + v_{2t}$$

حيث

$$\pi_2 = \frac{a_0 - b_0}{b_1 - a_1} \quad \pi_3 = \frac{a_2}{b_1 - a_1} \quad v_{2t} = \frac{u_{1t} - u_{2t}}{b_1 - a_1}$$

(ب) معادلات الشكل المختزل

$$Q_t = \pi_0 + \pi_1 Y_t + v_{1t}$$
$$P_t = \pi_2 + \pi_3 Y_t + v_{2t}$$

مهمة لأن Y_t غير مرتبطة مع v_{1t} ، v_{2t} ، وبالتالى يمكن الحصول على تقديرات متسقة لمعاملات الشكل المختزل ، π_0 ، π_1 ، π_2 و π_3 بتطبيق OLS على معادلات الشكل المختزل . وتعطى π_1 و π_3 على التوالى مجموع التأثيرات المباشرة وغير المباشرة للتغير فى Y على Q و P . إن التغير فى Y يسبب انتقال منحنى الطلب ، مما يؤثر على كل من P و Q التوازنية .

١٠ - ٤ بمعلومية نظام المعادلات الثلاث التالية ، (أ) اشرح لماذا لا يعتبر هذا نموذج معادلات آنية . (ب) هل يمكن استخدام *OLS* لتقدير كل معادلة فى هذا النظام ؟ لماذا ؟

$$Y_{1t} = a_0 + a_1X_t + u_{1t}$$
$$Y_{2t} = b_0 + b_1Y_{1t} + b_2X_t + u_{2t}$$
$$Y_{3t} = c_0 + c_1Y_{2t} + c_2X_t + u_{3t}$$

(أ) النظام السابق ليس آنياً لأنه بالرغم من أن Y_2 دالة فى Y_1 ، فإن Y_1 ليست دالة فى Y_2. وبالمثل ، Y_3 دالة فى Y_2 ، ولكن Y_2 ليست دالة فى Y_3. وبالتالى ، فإن خط السببية يجرى فى اتجاه واحد فقط وليس فى اتجاهين . فإذا تم تقدير Y_1 فى المعادلة الأولى ، فيمكن استخدام Y_1 (بالإضافة إلى X) لتقدير X_2 فى المعادلة الثانية . وبالمثل متى تم تقدير Y_2 فى المعادلة الثانية ، فيمكن استخدام Y_2 (بالإضافة إلى X) لتقدير Y_3 فى المعادلة الثالثة . وتسمى النماذج من هذا النوع بالنماذج المتواترة وليست آنية .

(ب) فى المعادلة الأولى ، المتغير الخارجى X غير مرتبط بحد الخطأ u_1 وبالتالى فإن OLS تعطى تقديرات معالم غير متحيزة للمعادلة الأولى . فى المعادلة الثانية ، X و Y غير مرتبطتين مع u_2 (أى أن Y_1 مرتبطة مع u_1 ولكن ليس مع u_2) ، وبالتالى تعطى OLS تقديرات معالم غير متحيزة للمعادلة الثانية . وينطبق نفس الشيء على المعادلة الثالثة . أى أنه يمكن تقدير النماذج المتواترة بالتطبيق المتتابع لطريقة OLS.

التمييز :

١٠ - ٥ (أ) ماذا يقصد بالتمييز ؟ (ب) متى تكون معادلة ما فى نظام مميزة بالضبط ؟ (جـ) زائدة التميز ؟ (د) ناقصة التمييز ؟ (هـ) هل هذه القواعد كافية للتمييز ؟

(أ) يشير التمييز إلى إمكانية أو عدم إمكانية الحصول على المعالم الهيكلية لنظام معادلات آنية من معالم الشكل المختزل . ويمكن أن تكون معادلة ما فى نظام مميزة بالضبط ، أو زائدة التمييز ، أو ناقصة التمييز . ويكون النظام ككل مميز بالضبط إذا كانت كل واحدة من معادلاته مميزة بالضبط .

(ب) تكون معادلة ما فى نظام مميزة بالضبط إذا كان عدد المتغيرات الخارجية المستبعدة من المعادلة مساوياً لعدد المتغيرات الداخلية فى المعادلة ناقصاً 1 . فى حالة المعادلة المميزة بالضبط ، يمكن حساب قيمة وحيدة للمعالم الهيكلية من معالم الشكل المختزل .

(جـ) تكون معادلة ما فى نظام زائدة التمييز إذا كان عدد المتغيرات الخارجية المستبعدة من المعادلة يتجاوز عدد المتغيرات الداخلية فى المعادلة ناقصاً 1 . فى حالة المعادلة زائدة التمييز ، يمكن حساب أكثر من قيمة عددية لبعض المعالم الهيكلية للمعادلة من معالم الشكل المختزل .

(د) تكون معادلة ما فى نظام ناقصة التمييز أو غير مميزة إذا كان عدد المتغيرات الخارجية المستبعدة من المعادلة أصغر من عدد المتغيرات الداخلية فى المعادلة ناقصاً 1 . وفى هذه الحالة لا يمكن حساب أى من المعالم الهيكلية من معالم الشكل المختزل .

(هـ) القواعد السابقة للتمييز (وتسمى شرط الترتيب) ضرورية وليست كافية . ولكن ، حيث أن هذه القواعد تعطى النتائج الصحيحة فى معظم الحالات ، فإنها الشروط الوحيدة المستخدمة فعلا هنا . الشرط الكافى للتمييز يعبر عنه شرط الرتبة ، والذى ينص على أنه فى نظام معادلات عددها G تكون معادلة معينة مميزة فقط إذا كان من الممكن الحصول على محدد واحد غير صفرى درجته $G - 1$ ، ذلك من معاملات المتغيرات المستبعدة من هذه المعادلة بالذات وإن كانت تدخل فى المعادلات الأخرى فى النموذج . وعندما يتوفر شرط الرتبة هذا ، فإن شرط الترتيب يتوفر تلقائياً . ولكن العكس غير صحيح .

١٠ – ٦ بمعلومية نموذج العرض – الطلب التالى (أ) حدد ما إذا كان الطلب و / أو العرض مميزا بالضبط ، زائد التمييز ، أو ناقص التمييز .

الطلب : $Q_t = a_0 + a_1 P_t + u_{1t} \quad a_1 < 0$

العرض : $Q_t = b_0 + b_1 P_t + u_{2t} \quad b_1 > 0$

(ب) ماذا يبين انحدار Q_t على P_t ؟

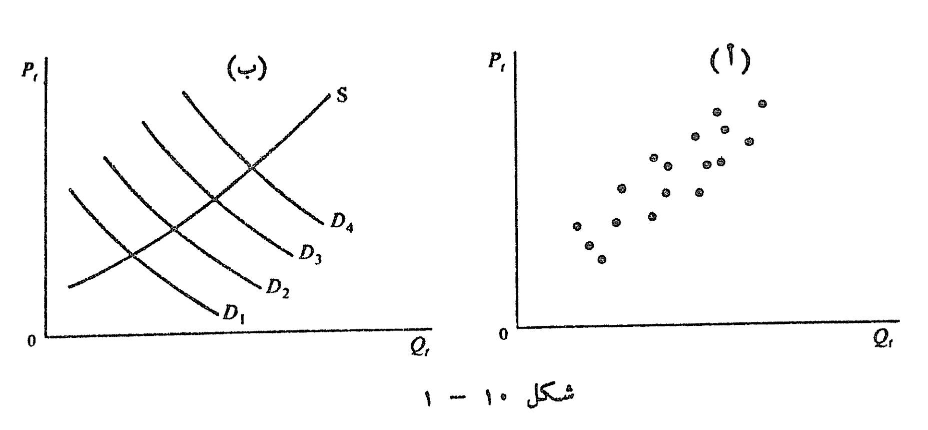
(أ) حيث أن نظام العرض – الطلب هذا لا يتضمن أى متغيرات خارجية ، فإن كلا من معادلتى العرض والطلب ناقصة التمييز . وفى هذه الحالة ، ليس هناك معادلات للشكل المختزل ، ولا يمكن حساب أى معاملات هيكلية . كل مشاهدة للسعر – الكمية تمثل كمية التوازن المشتراة والمباعة عند السعر المعين وتناظر تقاطع منحنى العرض ومنحنى الطلب (المجهولين) .

(ب) لا يعطى انحدار Q_t على P_t منحنى طلب أو منحنى عرض ، وإنما هجين من العرض والطلب ، والذى يجب الإشارة إليه ببساطة كخط انحدار .

١٠ – ٧ بالإشارة إلى نموذج العرض – الطلب فى المسألة ١٠ – ٢ (أ) حدد ما إذا كانت دالة الطلب و / أو العرض مميزة بالضبط ، زائدة التمييز ، أو ناقص التمييز (ب) اعط تفسيراً بيانياً لإجابتك فى (أ) (جـ) اشتق صيغة المعاملات الهيكلية من معاملات الشكل المختزل .

(أ) معادلة الطلب ناقصة التمييز لأنها لا تستبعد أى متغيرات خارجية . ولكن حيث أن هناك متغيراً خارجياً واحداً مستبعداً من معادلة العرض (أى Y) ومتغيرين داخليين فى المعادلة (أى Q و P) ، فان معادلة العرض مميزة بالضبط .

(ب) التغيرات فى Y تؤدى إلى نقلات فى منحنى الطلب مما يحدد منحنى العرض . يوضح شكل ١٠ – ١ (أ) شكل انتشار افتراضى للنقاط الناتجة عن التغيرات فى Y وفى حدود الخطأ ، بينما يوضح شكل ١٠ – ١ (ب) منحنى العرض الناتج الذى يمكن أن يتولد .



شكل ١٠ – ١

(جـ) يمكن حساب قيم وحيدة للمعاملات الهيكلية لمعادلة العرض (وهى المعادلة المميزة بالضبط) من معاملات الشكل المختزل فى المسألة ١٠ – ٣ كما يلى :

$$b_1 = \frac{\pi_1}{\pi_3} = \frac{\dfrac{b_1 a_2}{b_1 - a_1}}{\dfrac{a_2}{b_1 - a_1}}$$

$$b_0 = \pi_0 - b_1\pi_2 = \frac{a_0 b_1 - a_1 b_0}{b_1 - a_1} - \frac{b_1 a_0 + b_0 b_1}{b_1 - a_1} = \frac{b_0(b_1 - a_1)}{b_1 - a_1}$$

ولا يمكن اشتقاق صيغة المعاملات الهيكلية لمعادلة الطلب من معاملات الشكل المختزل لأن دالة الطلب فى هذا النموذج ناقصة التمييز .

١٠ – ٨ بالإشارة إلى نموذج الطلب – العرض المعطاة أدناه ، (أ) حدد ما إذا كانت دالة الطلب و / أو دالة العرض مميزة بالضبط ، زائدة التمييز ، أو ناقصة التمييز (ب) أوجد معادلات الشكل المختزل . (جـ) اشتق صيغة المعالم الهيكلية .

الطلب : $Q_t = a_0 + a_1 P_t + a_2 Y_t + u_{1t} \quad a_1 < 0, a_2 > 0$

العرض : $Q_t = b_0 + b_1 pP_t + b_2 T + u_{2t} \quad b_1 > 0, b_2 \lessgtr 0$

حيث $T =$ الاتجاه العام .

(أ) معادلة العرض مميزة بالضبط (كما فى المسألة ١٠ – ٧) لأنها تستبعد متغيراً خارجياً واحداً Y ، وتتضمن متغيرين ، P و Q . معادلة الطلب أصبحت الآن أيضاً مميزة بالضبط لأنها تستبعد متغيراً خارجياً واحداً T ، وتتضمن متغيرين داخليين P و Q .

(ب) يمكن الحصول على معادلات الشكل المختزل كما فى المسألة ١٠ – ٣ (أ) . وهى

$$Q_t = \left(\frac{a_0 b_1 - a_1 b_0}{b_1 - a_1}\right) + \left(\frac{a_2 b_1}{b_1 - a_1}\right) Y_t + \left(\frac{-a_1 b_2}{b_1 - a_1}\right) T + \left(\frac{b_1 u_{1t} - a_1 u_{2t}}{b_1 - a_1}\right)$$

$$P_t = \left(\frac{a_0 - b_0}{b_1 - a_1}\right) + \left(\frac{a_2}{b_1 - a_1}\right) Y_t + \left(\frac{-b_2}{b_1 - a_1}\right) T + \left(\frac{u_{1t} - u_{2t}}{b_1 - a_1}\right)$$

او

$$Q_t = \pi_0 + \pi_1 Y_t + \pi_2 T + v_{1t}$$
$$P_t = \pi_3 + \pi_4 Y_t + \pi_5 T + v_{2t}$$

حيث

$$\pi_0 = \frac{a_0 b_1 - a_a b_0}{b_1 - a_1} \quad \pi_1 = \frac{a_2 b_1}{b_1 - a_1} \quad \pi_2 = \frac{-a_1 b_2}{b_1 - a_1} \quad v_{1t} = \frac{b_1 u_{1t} - a_1 u_{2t}}{b_1 - a_1}$$

$$\pi_3 = \frac{a_0 - b_0}{b_1 - a_1} \quad \pi_4 = \frac{a_2}{b_1 - a_1} \quad \pi_5 = \frac{-b_2}{b_1 - a_1} \quad v_{2t} = \frac{u_{1t} - u_{2t}}{b_1 - a_1}$$

(جـ)

$$a_1 = \frac{\pi_2}{\pi_5} \quad \text{and} \quad b_1 = \frac{\pi_1}{\pi_4}$$

$$a_2 = \pi_4(b_1 - a_1) = \pi_4\left(\frac{\pi_1}{\pi_4} - \frac{\pi_2}{\pi_5}\right) \quad \text{and} \quad b_2 = -\pi_5(b_1 - a_1) = \pi_5\left(\frac{\pi_2}{\pi_5} - \frac{\pi_1}{\pi_4}\right)$$

$$a_0 = \pi_3(b_1 - a_1) + b_0 = \pi_3\left(\frac{\pi_0}{\pi_3} - \frac{\pi_2}{\pi_5}\right) \quad \text{and} \quad b_0 = \pi_3(b_1 - a_1) + a_0 = \pi_3\left(\frac{\pi_0}{\pi_3} - \frac{\pi_1}{\pi_4}\right)$$

١٠ — ٩ بالإشارة إلى نموذج العرض — الطلب المعطى من قبل (أ) حدد ما إذا كانت معادلة الطلب و / أو العرض مميزة بالضبط ، زائدة التمييز ، أو ناقصة التمييز . (ب) احسب معالم الميل الهيكلية .

الطلب : $Q_t = a_0 + a_1 P_t + a_2 Y_t + a_3 W_t + u_{1t}$

العرض : $Q_t = b_0 + b_1 P_t + u_{2t}$

حيث W_t هى الثروة والتوقع أن تكون $a_3 > 0$.

(أ) معادلة الطلب ناقصة التمييز لأنها لا تستبعد أى متغيرات خارجية . ولكن حيث أن هناك متغيرين خارجيين مستبعدين من معادلة العرض (أى ، Y و W) ومتغيرين داخليين تتضمنهما المعادلة (أى ، Q و P) ، فان معادلة العرض زائدة التمييز .

(ب) لحساب معالم الميل الهيكلية ، يجب إيجاد معادلات الشكل المختزل . ويتم الحصول عليها كما فى المسألة ١٠ — ٧ (جـ) وهى

$$Q_t = \pi_0 + \pi_1 Y_t + \pi_2 W_t + v_{1t}$$
$$P_t = \pi_3 + \pi_4 Y_t + \pi_5 W_t + v_{2t}$$

حيث

$$\pi_0 = \frac{a_0 b_1 - a_1 b_0}{b_1 - a_1} \quad \pi_1 = \frac{a_2 b_1}{b_1 - a_1} \quad \pi_2 = \frac{a_3 b_1}{b_1 - a_1}$$
$$\pi_3 = \frac{a_0 - b_0}{b_1 - a_1} \quad \pi_4 = \frac{a_2}{b_1 - a_1} \quad \pi_5 = \frac{a_3}{b_1 - a_1}$$

ويمكن حساب b_1 من

$$\frac{\pi_1}{\pi_4} = b_1 \quad \text{or} \quad \frac{\pi_2}{\pi_5} = b_1$$

وهذان التقديران للمعلمة b_1 سيكونان عادة مختلفين ، مما يعكس حقيقة أن معادلة العرض هى الآن زائدة التمييز . وكما فى المسألة ١٠ — ٧ (جـ) ، لا يمكن حساب المعاملات الهيكلية لدالة الطلب من معاملات الشكل المختزل لأن دالة الطلب فى هذا النموذج ناقصة التمييز .

التقدير : المربعات الصغرى غير المباشرة :

١٠ — ١٠ (أ) متى يمكن استخدام المربعات الصغرى غير المباشرة ؟ (ب) ماذا تتضمن ؟ (جـ) ما هى بعض نواحى القصور لاستخدام المربعات الصغرى غير المباشرة ؟

(أ) المربعات الصغرى غير المباشرة (ILS) هى طريقة لحساب قيم معالم هيكلية متسقة للمعادلات المميزة بالضبط فى نظام من المعادلات الآنية .

(ب) تتضمن ILS استخدام OLS لتقديرات الشكل المختزل للنظام ثم استخدام المعالم المقدرة للشكل المختزل لحساب تقديرات معالم هيكلية وحيدة ومتسقة ، كما سبق الإشارة فى المسائل ١٠ — ٧ (جـ) ، ١٠ — ٨ (جـ) ، و ١٠ — ٩ (ب) .

(جـ) من عيوب استخدام ILS أنها لا تعطى الخطأ المعيارى للمعالم الهيكلية المحسوبة ، وعملية حسابها معقدة إلى حد كبير (وخارج نطاق هذا الكتاب) . وعيب آخر لطريقة ILS أنه لا يمكن استخدامها لحساب تقديرات معالم هيكلية وحيدة ومتسقة من معاملات الشكل المختزل للمعادلات زائدة التمييز لنموذج المعادلات الآنية .

١٠ – ١١ يعطى جدول ١٠ – ٢ رقما قياسياً لإنتاج المحاصيل Q ، أسعار المحاصيل P ، ومتوسط دخل الفرد المتاح Y ، بأسعار ١٩٧٢ ، فى الولايات المتحدة ١٩٥٠ – ١٩٧٩ . افترض أنه بنهاية كل عام يباع كل المعروض أى أن Q_t تمثل كلا من الكمية المشتراة والمباعة فى السنة t . قدر باستخدام OLS معادلات الشكل المختزل المعطاة فى مسألة ١٠ – ٣ (أ) ، (ب) احسب المعالم الهيكلية للعرض من معاملات الشكل المختزل (جـ) كيف تقارن هذه مع المعالم الهيكلية التى يحصل عليها باجراء انحدار Q_t على P_t مباشرة ؟

جدول ١٠ – ٢ الرقم القياسى لإنتاج المحاصيل ، الأسعار ودخل الفرد المتاح بأسعار ١٩٧٢ : الولايات المتحدة ، ١٩٥٠ – ١٩٧٩

السنة	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964
Q	76	78	81	81	79	82	82	80	89	89	93	91	92	96	93
P	103	118	119	107	108	103	104	100	99	98	99	101	103	107	106
Y	2,386	2,408	2,434	2,491	2,476	2,517	2,643	2,650	2,636	2,696	2,697	2,725	2,796	2,849	3,009
السنة	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
Q	99	95	100	103	104	100	112	113	119	110	121	121	130	131	144
P	103	106	100	100	97	100	108	114	175	224	201	197	192	204	223
Y	3,152	3,274	3,371	3,464	3,515	3,619	3,714	3,837	4,062	3,973	4,025	4,144	4,285	4,449	4,509

المصدر : التقرير الاقتصادى للرئيس ، مكتب حكومة الولايات المتحدة للطباعة ، واشنطن ، ١٩٨٠ صفحة ٣١٠ ، ٣١٢ ، ٣١٢ .

(أ) معادلات الشكل المختزل المقدرة (من المسألة ١٠ – ٣ (أ)) هى

$$\hat{Q}_t = \underset{(4.47)}{18.0313} + \underset{(20.60)}{0.0252}\,Y_t \qquad R^2 = 0.94$$

$$\hat{P}_t = \underset{(-1.21)}{-29.9304} + \underset{(6.50)}{0.0487}\,Y_t \qquad R^2 = 0.60$$

(ب)

$$\hat{b}_1 = \hat{\pi}_1 = \frac{0.0252}{0.0487} = 0.5175 \qquad \text{[see Prob. 10.7(c)]}$$

$$\hat{b}_0 = \hat{\pi}_0 - \hat{b}_1\hat{\pi}_2 = 18.0313 - 0.5174(-29.9304) = 33.5173$$

$\hat{b}_0$ و $\hat{b}_1$ هى مقدرات متسقة للمعالم b_0 و b_1 على الترتيب ، والمعادلة الهيكلية للطلب (المقدرة باستخدام ILS) هى

$$\hat{Q}_t = 33.53 + 0.52P_t$$

(جـ) باجراء انحدار Q_t على P_t مباشرة نحصل على :

$$\hat{Q}_t = \underset{(8.92)}{57.98} + \underset{(6.73)}{0.33}\,P_t \qquad R^2 = 0.62$$

قيم $\hat{b}_0$ و $\hat{b}_1$ التى تم الحصول عليها باجراء انحدار Q_t على P_t هى تقديرات متحيزة وغير متسقة لمعالم العرض .

١٠ – ١٢ بالإشارة إلى نموذج الطلب – العرض فى المسألة ١٠ – ٨ وباستخدام بيانات جدول ١٠ – ٢ وقيم الاتجاه العام $T = 1,2,3,\ldots,30$ ، (أ) احسب معالم هيكلية متسقة لمعادلة الطلب . (ب) كيف تقارن هذه مع المعالم الهيكلية التى يتم الحصول عليها بتقدير معادلة الطلب مباشرة باستخدام OLS ؟

(أ) حيث أن معادلة الطلب مميزة بالضبط (أنظر المسألة ١٠ – ٨ (أ)) فإنه يمكننا استخدام ISL للحصول على قيم معالم هيكلية متسقة للطلب . معادلات الشكل المختزل المقدرة (من المسألة ١٠ – ٨ (ب)) هى :

$$\hat{Q}_t = \underset{(2.35)}{28.7649} + \underset{(3.34)}{0.0198Y_t} + \underset{(0.93)}{0.4292T} \qquad R^2 = 0.94$$

$$\hat{P}_t = \underset{(-2.78)}{-191.1760} + \underset{(3.89)}{0.1296Y_t} - \underset{(-2.48)}{6.4475T} \qquad R^2 = 0.68$$

حيث

$$\pi_0 = 28.7649, \quad \pi_1 = 0.0198, \quad \pi_2 = 0.4292,$$
$$\pi_3 = -191.1760, \quad \pi_4 = 0.1296, \quad \pi_5 = -6.4475$$

باستخدام المعادلات المعطاة فى المسألة ١٠ – ٨ (جـ) ، نحصل على :

$$\hat{a}_1 = \frac{\hat{\pi}_2}{\hat{\pi}_5} = \frac{0.4292}{-6.4475} = -0.0666$$

$$\hat{a}_2 = \hat{\pi}_4\left(\frac{\hat{\pi}_1}{\hat{\pi}_4} - \frac{\hat{\pi}_2}{\hat{\pi}_5}\right) = 0.1296\left(\frac{0.0198}{0.1296} - \frac{0.4292}{-6.4475}\right) = 0.0284$$

$$\hat{a}_0 = \hat{\pi}_3\left(\frac{\hat{\pi}_0}{\hat{\pi}_3} + \frac{\hat{\pi}_2}{\hat{\pi}_5}\right) = -191.1760\left(\frac{28.7649}{-191.1760} + \frac{0.4292}{6.4475}\right) = 16.0386$$

وعليه فان معادلة الطلب المقدرة باستخدام ILS (والتى تظهر تقديرات معالم متسقة) هى :

$$\hat{Q}_t = 16.04 - 0.07P_t + 0.03Y_t$$

(ب) تقدير OLS لدالة الطلب هى :

$$\hat{Q}_t = \underset{(4.65)}{19.12} + \underset{(1.19)}{0.04P_t} + \underset{(12.18)}{0.02Y_t} \qquad R^2 = 0.94$$

قيم $\hat{a}_0$ ، $\hat{a}_1$ و $\hat{a}_2$ متحيزة وغير متسقة . والحقيقة ، أن $\hat{a}_1$ تأخذ الإشارة الخطأ (ولكنها ليست معنوية إحصائياً) .

التقدير : المربعات الصغرى على مرحلتين :

١٠ – ١٣ (أ) متى يمكن استخدام 2SLS ؟ (ب) ماذا تتضمن ؟ (جـ) ما هى مزايا 2SLS بالنسبة إلى ILS ؟

(أ) المربعات الصغرى على مرحلتين 2SLS هى طريقة لتقدير قيم معالم هيكلية متسقة للمعادلات المميزة بالضبط أو زائدة التمييز لنظام معادلات آنية وبالنسبة للمعادلات المميزة بالضبط ، تعطى 2SLS نفس نتيجة ILS .

(ب) يتضمن تقدير 2SLS تطبيق OLS على مرحلتين . فى المرحلة الأولى ، يتم إجراء انحدار كل متغير داخلى على كل المتغيرات المحددة سلفاً فى النظام . هذه الآن هى معادلات الشكل المختزل . فى المرحلة الثانية ، تستخدم قيم المتغيرات الداخلية المقدرة بدلا من الفعلية لتقدير المعادلات الهيكلية للنموذج . ويتم الحصول على القيم المقدرة للمتغيرات الداخلية بالتعويض بالقيم الفعلية للمتغيرات الخارجية فى معادلات الشكل المختزل . القيم المقدرة للمتغيرات الداخلية غير مرتبطة مع حدود الخطأ مؤدية بذلك إلى تقديرات 2SLS متسقة للمعالم الهيكلية .

(جـ) من مزايا 2SLS على ILS أنه يمكن استخدام 2SLS لحصول على تقديرات معالم هيكلية متسقة للمعادلات زائدة التميز كما بالنسبة للمعادلات المميزة بالضبط فى نظام معادلات آنية . والميزة الهامة الثانية أن 2SLS (ولكن ليس ILS) تعطى الخطأ المعيارى للمعالم الهيكلية المقدرة مباشرة . وحيث أن معظم النماذج المميزة هى فى الواقع زائدة التمييز ، فان 2SLS مفيدة جداً . وبالتأكيد ، تعتبر 2SLS أبسط ، وواحدة من أفضل طرق تقدير المعادلات الآنية وأكثرها شيوعاً .

١٠ – ١٤ بالنسبة لنموذج الطلب – العرض فى المسألة ١٠ – ٨ وباستخدام بيانات جدول ١٠ – ٢ لتقدير معادلة الطلب ، (أ) بين نتائج المرحلة الأولى لتقدير 2SLS (ب) بين نتائج المرحلة الثانية لتقدير 2SLS . (جـ) كيف تقارن هذه النتائج مع تقدير ILS لمعادلة الطلب السابق إيجادها فى المسألة ١٠ – ١٢ (أ) ؟

(أ) نتائج المرحلة الأولى لتقدير 2SLS لمعادلة الطلب هى :

$$\hat{P}_t = \underset{(-2.78)}{191.1760} + \underset{(3.89)}{0.1296\,Y_t} - \underset{(-2.48)}{6.4475\,T} \qquad R^2 = 0.68$$

(ب) نتائج المرحلة الثانية لتقدير 2SLS لمعادلة الطلب هى :

$$\hat{Q}_t = \underset{(3.50)}{16.04} - \underset{(-0.93)}{0.07\,P_t} + \underset{(7.69)}{0.03\,Y_t} \qquad R^2 = 0.94$$

(جـ) حيث أن معادلة الطلب فى المسألة ١٠ – ٨ مميزة بالضبط ، فإن تقدير 2SLS يعطى نتائج تطابق تقدير ILS (أنظر المسألة ١٠ – ١٢ (أ) . ولكن باستخدام 2SLS (مقارنة مع ILS) ، فإننا نحصل أيضاً على الأخطاء المعيارية للمعالم الهيكلية المقدرة مباشرة . لاحظ أن P_t ليست معنوية إحصائياً ، مما يعكس حقيقة أن مشاهدات السعر – الكمية فى أسواق المحاصيل أكثر ملاءمة لتقدير معادلة العرض العملية عن معادلة الطلب .

١٠ – ١٥ يتضمن جدول ١٠ – ٣ المتغير الإضافى الثروة ، W ، مقيساً هنا باجمالى الأصول المتداولة ، بالبليون دولار بالإضافة إلى بيانات جدول ١٠ – ٢ للولايات المتحدة للسنوات ١٩٥٢ – ١٩٧٩ بالنسبة لنموذج الطلب – العرض فى المسألة ١٠ – ٩ ، قدر معادلة العرض باستخدام (أ) 2SLS ، (ب) OLS .

جدول ١٠ – ٣ الرقم القياسى لإنتاج المحاصيل ، الأسعار ، الدخل المتاح للفرد بأسعار ١٩٧٢ ، وإجمالى الأصول المتداولة : للولايات المتحدة ، ١٩٥٢ – ١٩٧٩ .

السنة	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961
Q	81	81	79	82	82	80	89	89	93	91
P	119	107	108	103	104	100	99	98	99	101
Y	2,434	2,491	2,476	2,517	2,643	2,650	2,636	2,696	2,697	2,725
W	269.1	284.6	295.3	314.8	325.4	338.0	354.4	373.3	386.8	410.7
السنة	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971
Q	92	96	93	99	95	100	103	104	100	112
P	103	107	106	103	106	100	100	97	100	108
Y	2,796	2,849	3,009	3,152	3,274	3,371	3,464	3,515	3,619	3,714
W	442.1	479.3	515.5	559.6	587.3	638.3	696.8	722.7	769.8	854.9
السنة	1972	1973	1974	1975	1976	1977	1978	1979		
Q	113	119	110	121	121	130	131	144		
P	114	175	224	201	197	192	204	223		
Y	3,837	4,062	3,973	4,025	4,144	4,285	4,449	4,509		
W	966.8	1,086.1	1,174.2	1,295.6	1,428.4	1,598.7	1,775.3	1,957.7		

المصدر : التقرير الاقتصادى للرئيس ، ومكتب حكومة الولايات المتحدة للطباعة ، واشنطن ، ١٩٨٠ صفحة ٣١٠ ، ٣١٢ ، ٣٢٩ ، ٢٧٣ .

(أ) حيث أن معادلة العرض فى المسألة ١٠ – ٩ زائدة التمييز ، فإن 2SLS تكون أسلوب تقدير ملائم للحصول على معالم هيكلية متسقة . المرحلة الأولى هى

$$\hat{P}_t = \underset{(3.80)}{170.90} - \underset{(-2.31)}{0.04 Y_t} + \underset{(5.24)}{0.14 W_t} \qquad R^2 = 0.82$$

المرحلة الثانية هى :

$$\hat{Q}_t = \underset{(6.85)}{49.69} + \underset{(7.42)}{0.40 \hat{P}_t} \qquad R^2 = 0.60$$

تقدير OLS (غير الملائم) لمعادلة العرض هو :

$$\hat{Q}_t = \underset{(9.74)}{60.62} + \underset{(6.86)}{0.31 P_t} \qquad R^2 = 0.64$$

تقديرات معالم العرض متسقة باستخدام 2SLS وغير متسقة باستخدام OLS . كلما ارتفعت R^2 فى المرحلة الأولى لتقدير 2SLS ، كلما اقتربت تقديرات معالم 2SLS و OLS .

مسائل إضافية

نماذج المعادلات الآنية :

١٠ – ١٦ تمثل المعادلتان التاليتان نموذج أجور – اسعار بسيط :

$$W_t = a_0 + a_1 P_t + a_2 Q_t + u_{1t}$$
$$P_t = b_0 + b_1 W_t + u_{2t}$$

حيث W_t هو الأجور فى الفترة t و P تمثل الأسعار و Q تمثل الإنتاجية . (أ) لماذا يعتبر هذا نموذج معادلات آنية ؟ (ب) ما هى المتغيرات الداخلية والخارجية ؟ (ج) لماذا يعطى تقدير معادلات W و P باستخدام OLS تقديرات معالم متحيزة وغير متسقة ؟

الإجابة (أ) هذا النموذج ذو المعادلتين له طبيعة آنية لأن W دالة فى P و P دالة فى W ، وبالتالى فان W و P تتحددان معاً . (ب) المتغيرات الداخلية هى W و P . المتغير الخارجى هو Q . (ج) تقدير دالة W باستخدام OLS يعطى تقديرات معالم متحيزة وغير متسقة لأن P ترتبط مع u_1 . وبالمثل ، فان تقدير المعادلة الثانية ، معادلة P باستخدام OLS يعطى تقديرات معالم متحيزة وغير متسقة لأن W مرتبطة مع u_2 .

١٠ – ١٧ (أ) أوجد معادلات الشكل المختزل للنموذج فى المسألة ١٠ – ١٦ . (ب) لماذا هى هامة ؟ (ج) ماذا تقيس معاملات الشكل المختزل فى هذا النموذج الكلى ؟

الإجابة : (أ)

$$W_t = \frac{a_0 + a_1 b_0}{1 - a_1 b_1} + \frac{a_2}{1 - a_1 b_1} Q_t + \frac{u_{1t} + a_1 u_{2t}}{1 - a_1 b_1} \quad \text{أو} \quad W_t = \pi_0 + \pi_1 Q_t + v_{1t}$$

$$P_t = \frac{b_0 + a_0 b_1}{1 - a_1 b_1} + \frac{a_2 b_1}{1 - a_1 b_1} Q_t + \frac{b_1 u_{1t} + u_{2t}}{1 - a_1 b_1} \quad \text{أو} \quad P_t = \pi_2 + \pi_3 Q_t + v_{2t}$$

(ب) معادلات الشكل المختزل مهمة لأنها تعبر عن كل متغير داخلى فى النموذج كدالة فى المتغيرات الخارجية فقط ، وبالتالى تعطى OLS تقديرات معالم متسقة . (ج) تعطى معالم الشكل المختزل إجمالى التأثير المباشر وغير المباشر للتغير فى اى متغير خارجى فى النموذج على كل متغير داخلى فى النموذج .

١٠ – ١٨ (أ) ما نوع النموذج الآتى ؟ (ب) كيف يمكن تقدير معادلات هذا النموذج ؟

$$Y_{1t} = a_0 + a_1 X_{1t} + u_{1t}$$
$$Y_{2t} = b_0 + b_1 Y_{1t} + b_2 X_{2t} + u_{2t}$$
$$Y_{3t} = c_0 + c_1 Y_{1t} + c_2 Y_{2t} + c_3 X_{2t} + u_{3t}$$

الإجابة : (أ) النموذج متواتر (ب) يمكن تقدير معادلات النموذج بتطبيق OLS بالتتابع ، بدءاً بالمعادلة الأولى .

التمييز :

١٠ – ١٩ لو لم يشمل النموذج الكلى البسيط فى المسألة ١٠ – ١٦ المتغير Q_t ، (أ) هل تكون المعادلة الأولى مميزة بالضبط ، زائدة التمييز ، أو ناقصة التمييز ؟ (ب) ماذا عن المعادلة الثانية ؟

الإجابة : (أ) تكون المعادلة الأولى ناقصة التمييز . (ب) تكون المعادلة الثانية ناقصة التمييز أيضاً .

١٠ — ٢٠ بالنسبة للنموذج الكلى فى المسألة ١٠–١٦ حدد (أ) ما إذا كانت المعادلة الأولى متميزة بالضبط ، زائدة التمييز ، أو ناقصة التمييز ؟ (ب) ماذا عن المعادلة الثانية ؟ (جـ) ما هى قيم المعالم الهيكلية ؟

الإجابة : (أ) المعادلة الأولى ناقصة التمييز (ب) المعادلة الثانية مميزة بالضبط. (جـ)$b_1 = \pi_3/\pi_1$ ، $b_0 = \pi_2 - b_1\pi_0$
لا يمكن حساب a_1 و a_2 من معاملات الشكل المختزل لأن معادلة W ناقصة التمييز .

١٠ — ٢١ لو احتوت المعادلة الثانية فى النموذج الكلى فى المسألة ١٠–١٦ على متغير إضافى Y (GNP) ، (أ) حدد ما إذا كانت معادلة W و / أو P مميزة بالضبط ، زائدة التمييز ، أو ناقصة التمييز (ب) أوجد معادلات الشكل المختزل . (جـ) اشتق صيغ المعالم الهيكلية .

الإجابة : (أ) كلاً من المعادلة الأولى ، معادلة W ، والمعادلة الثانية ، معادلة P ، أصبحت الآن مميزة بالضبط .

(ب)

$$W_t = \frac{a_0 + a_1b_0}{1 - a_1b_1} + \frac{a_2}{1 - a_1b_1} Q_t + \frac{a_1b_2}{1 - a_1b_1} Y_t + \frac{u_{1t} + a_1u_{2t}}{1 - a_1b_1}$$

$$P_t = \frac{a_0b_1 + b_0}{1 - a_1b_1} + \frac{a_2b_1}{1 - a_1b_1} Q_t + \frac{b_2}{1 - a_1b_1} Y_t + \frac{b_1u_{1t} + u_{2t}}{1 - a_1b_1}$$

أو

$$W_t = \pi_0 + \pi_1 Q_t + \pi_2 Y_t + v_{1t}$$
$$P_t = \pi_3 + \pi_4 Q_t + \pi_5 Y_t + v_{2t}$$

(جـ)

$$a_1 = \frac{\pi_2}{\pi_5} \quad \text{و} \quad b_1 = \frac{\pi_4}{\pi_1}$$

$$a_2 = \pi_2\left(\frac{\pi_1}{\pi_2} - \frac{\pi_4}{\pi_5}\right) \quad \text{و} \quad b_2 = \pi_2\left(\frac{\pi_5}{\pi_2} - \frac{\pi_4}{\pi_1}\right)$$

$$a_0 = \pi_3\left(\frac{\pi_0}{\pi_3} - \frac{\pi_2}{\pi_5}\right) \quad \text{و} \quad b_0 = \pi_0\left(\frac{\pi_3}{\pi_0} - \frac{\pi_4}{\pi_1}\right)$$

١٠ — ٢٢ لو تضمنت المعادلة الأولى فى المسألة ١٠ – ١٦ المتغير الإضافى P_{t-1} (السعر مبطأ سنة واحدة) ، (أ) هل تصبح المعادلات مميزة بالضبط ، أو زائدة التمييز أو ناقصة التمييز ؟ (ب) ما قيمة معالم الميل الهيكلية ؟

الإجابة : (أ) المعادلة الأولى ، معادلة W ، ناقصة التمييز ، بينما المعادلة الثانية ، معادلة P ، وزائدة التمييز .
(ب) $b_1 = \pi_4/\pi_1$ أو π_5/π_2 ، مما يعكس حقيقة أن معادلة P أصبحت الآن زائدة التمييز ؛ لا يمكن حساب a_1 ، a_2 وa_3 لأن معادلة W ناقصة التمييز .

التقدير : المربعات الصغرى غير المباشرة :

١٠ — ٢٣ يعطى جدول ١٠ – ٤ متوسط أجر الساعة الإجمالى فى القطاع الخاص غير الزراعى W ، الرقم القياسى لأسعار المستهلكين P ، الإنتاج / ساعة فى قطاع الأعمال غير الزراعى Q وGNP ، Y ، فى الولايات المتحدة من ١٩٦٠ إلى ١٩٧٩ .
(أ) قدر معادلات الشكل المختزل فى المسألة ١٠ – ١٧ (أ) . (ب) احسب المعاملات الهيكلية لمعادلة P من معاملات الشكل المختزل (جـ) كيف تقارن هذه مع المعالم الهيكلية التى يمكن الحصول عليها باجراء انحدار P على W مباشرة ؟

جدول ١٠ – ٤ الأجور ، الرقم القياسى للأسعار ، الإنتاجية ، و GNP : الولايات المتحدة ، ١٩٦٠ – ١٩٧٩ .

السنة	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969
W	2.09	2.14	2.22	2.28	2.36	2.46	2.56	2.68	2.85	3.04
P	88.7	89.6	90.6	91.7	92.9	94.5	97.2	100.0	104.2	109.8
Q	80.9	83.0	86.6	89.6	92.8	95.9	98.4	100.0	103.2	102.9
Y	506.0	523.3	563.8	594.7	635.7	688.1	753.0	796.3	868.5	935.5

السنة	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979
W	3.23	3.45	3.70	3.94	4.24	4.53	4.86	5.25	5.69	6.16
P	116.3	121.3	125.3	133.1	147.7	161.2	170.5	181.5	195.4	217.4
Q	103.0	106.2	110.1	112.0	108.5	110.5	114.4	116.2	116.8	115.5
Y	982.4	1,063.4	1,171.1	1,306.6	1,412.9	1,528.8	1,702.2	1,899.5	2,127.6	2,368.5

المصدر : التقرير الإقتصادى للرئيس ، مكتب حكومة الولايات المتحدة للطباعة ، واشنطن ، ١٩٨٠ ، صفحة ٢٤٤ ، ٢٥٩ ، ٢٤٦ ، ٢٠٣ .

الإجابة : (أ)

$$\hat{W}_t = -6.7961 + 0.1005 Q_t \qquad R^2 = 0.80$$
$$(-5.56) \qquad (8.45)$$

$$\hat{P}_t = -178.8330 + 2.9834 Q_t \qquad R^2 = 0.72$$
$$(-3.96) \qquad (6.80)$$

(ب) $\hat{b}_1 = 29.6856;\ \hat{b}_0 = 22.9133$ (جـ) باستخدام OLS $\hat{b}_1 = 31.1175$ و $\hat{b}_0 = 17.9538$

١٠ – ٢٤ بالنسبة للنموذج فى المسألة ١٠–٢١ ، (أ) قدر معادلات الشكل المختزل ، (جـ) احسب المعاملات الهيكلية لمعادلة W من معاملات الشكل المختزل . (جـ) كيف تقارن هذه مع المعاملات الهيكلية لمعادلة W التى يمكن الحصول عليها باستخدام OLS؟

الإجابه : (أ)

$$\hat{W}_t = 0.5466 + 0.0050 Q_t + 0.0022 Y_t \qquad R^2 = 0.997$$
$$(2.08) \qquad (1.57) \qquad (34.20)$$

$$\hat{P}_t = 92.5222 - 0.5473 Q_t + 0.0802 Y_t \qquad R^2 = .0996$$
$$(10.18) \qquad (-5.01) \qquad (36.49)$$

(ب) $\hat{a}_0 = -1.9570,\ \hat{a}_1 = 0.0271$ و $\hat{a}_2 = 0.0200$ (جـ) باستخدام OLS $\hat{a}_0 = -1.9706,\ \hat{a}_1 = 0.0270$ و $\hat{a}_2 = 0.0198$

١٠ – ٢٥ بالنسبة للنموذج فى المسألة ١٠ – ٢١ ، أكتب المعادلة الهيكلية لمعادلة P المقدرة باستخدام (أ) ILS ، (ب) OLS .

الإجابة : (أ)

$$\hat{P}_t = 152.9570 - 110.5730 W_t + 0.3201 Y_t$$

(ب)

$$\hat{P}_t = 39.7567 + 8.0649 W_t + 0.0522 Y_t \qquad R^2 = 0.991$$
$$(3.37) \qquad (0.66) \qquad (1.89)$$

المربعات الصغرى على مرحلتين :

١٠ – ٢٦ بالنسبة للنموذج فى المسألة ١٠ – ٢١ وباستخدام بيانات جدول ١٠ – ٤ لتقدير معادلة W ، (أ) بين نتائج المرحلة الأولى لتقدير 2SLS ، (ب) بين نتائج المرحلة الثانية لتقدير 2SLS (جـ) كيف تقارن هذه النتائج مع تقدير ILS لمعادلة W السابق إيجادها فى مسألة ١٠ – ٢٤ ؟

الإجابة : (أ) $\hat{P}_t = 95.5222 - 0.5473\,Q_t + 0.0802\,Y_t \qquad R^2 = 0.996$

$\qquad (10.18) \quad (-5.01) \quad (36.49)$

(ب) $\hat{W}_t = -1.9570 + 0.0271\,\hat{P}_t + 0.0198\,Q_t \qquad R^2 = 0.997$

$\qquad (-9.47) \quad (34.20) \quad (7.11)$

(جـ) تتطابق النتائج ، لكن مع تقدير 2SLS ، نحصل أيضاً على الأخطاء المعيارية . المعالم الهيكلية المقدرة باستخدام 2SLS و ILS متسقة .

١٠ – ٢٧ بالنسبة للنموذج فى المسألة ١٠ – ٢٢ وبيانات جدول ١٠ – ٤ ، قدر معادلة P باستخدام (أ) 2SLS ، (ب) OLS. (جـ) هل تقديرات المعالم الهيكلية فى (أ) و (ب) غير متحيزة ؟ متسقة ؟

الإجابة : (أ) $\hat{P}_t = 16.50 + 31.44\,\hat{W}_t \qquad R^2 = 0.988$

$\qquad (5.28) \quad (37.89)$

(ب) $\hat{P}_t = 16.59 + 31.42\,W_t \qquad R^2 = 0.990$

$\qquad (5.89) \quad (41.93)$

(جـ) تقديرات المعالم فى (أ) و (ب) متحيزة ، ولكنها متسقة فى (أ) وغير متسقة فى (ب) .

امتحان اقتصاد قياسى

١ – يعطى جدول (١) كية المعروض من سلعة ما ، Y ، عند أسعار مختلفة ، X ، مع بقاء الأشياء الأخرى ثابتة .

(أ) قدر معادلة انحدار Y على X . (ب) اختبر المعنوية الإحصائية لتقديرات المعالم عند مستوى معنوية 5% (ج) أوجد R^2 وضع كل النتائج السابقة فى شكل نمطى مختزل . (د) تنبأ بقيمة $\hat{Y}$ واحسب فترة ثقة أو تنبؤ 95% عند $X = 10$

جدول (١) الكمية المعروضة عند مختلف الأسعار

n	1	2	3	4	5	6	7	8
Y	12	14	10	13	17	12	11	15
X	5	11	7	8	11	7	6	9

٢ – افترض أنه باستخدام 24 مشاهدة سنوية عن الكميات المطلوبة بالكيلوجرام من سلعة ما سنوياً ، Y ، وسعرها بالدولارات ، X_1 ودخول المستهلكين بالألف دولار ، X_2 ، وسعر سلعة بديلة بالدولار ، X_3 وتم الحصول على الانحدار المقدر التالى ، حيث الأرقام داخل الأقواس هى الأخطاء المعيارية :

$$\hat{Y} = 13 - 7X_1 + 2.4X_2 - 4X_3$$
$$(2) \quad (0.8) \quad (18)$$

(ا) حدد ما إذا كانت إشارات المعالم تؤكد ماهو متوقع طبقاً لنظرية الطلب .

(ب) هل معالم الميل المقدرة معنوية عند مستوى 5%؟ (ج) أوجد R^2 ، إذا كانت $\Sigma y^2=40$ ، $\Sigma yx_1 = 10$ و $\Sigma yx_2 = 45$ (حيث تشير الحروف الصغيرة إلى الانحرافات عن المتوسطات ، (د) أوجد $\bar{R}^2$ (ه) هل R^2 تختلف معنوياً عن الصفر عند مستوى 5% ؟ (و) أوجد الخطأ المعيارى للانحدار (ز) أوجد معاملات المرونة السعرية والدخلية للطلب عند المتوسطات إذا كان $\bar{Y} = 32$ ، $\bar{X}_1 = 8$ ، $\bar{X}_2 = 16$.

٣ – عندما يجرى انحدار مستوى إنفاق قطاع الأعمال على التجهيزات الجديدة للشركات غير الصناعية فى الولايات المتحدة Y_t من ١٩٦٠ إلى ١٩٧٩ على GNP ، X_{1t} ، والرقم القياسى لأسعار المستهلكين ، X_{2t} ، يتم الحصول على النتائج التالية

$$\hat{Y}_t = 31.75 + 0.08X_{1t} - 0.58X_{2t} \qquad R^2 = 0.98$$
$$(6.08) \quad (-3.08) \qquad d = 0.77$$

(أ) كيف يمكن ان تعرف ما إذا كان هناك ارتباط ذاتى ؟ ماذا يقصد بالارتباط الذاتى ؟ لماذا يمثل الارتباط الذاتى مشكلة ؟

(ب) كيف يمكن تقدير ρ ، معامل الارتباط الذاتى ؟ (ج) كيف يمكن استخدام قيمة ρ لتحويل المتغيرات لتصحيح الارتباط الذاتى ؟ كيف يمكن إيجاد القيمة الأولى للمتغيرات المحولة ؟ (د) هل هناك أى دليل على وجود ارتباط ذاتى باق من النتائج التالية التى تم الحصول عليها من إجراء الانحدار على المتغيرات المحولة (مشارا إليها بنجمة) ؟

$$Y_t^* = 3.79 + 0.04X_{1t}^* - 0.05X_{2t}^* \qquad R^2 = 0.96$$
$$(8.10) \quad (-0.72) \qquad d = 0.89$$

ماذا يمكن ان يكون السبب فى اى ارتباط ذاتى باق ؟ كيف يمكن تصحيح هذا ؟

٤ – تمثل المعادلتان التاليتان نموذجاً كلياً بسيطاً :

$$R_t = a_0 + a_1 M_t + a_2 Y_t + u_{1t}$$
$$Y_t = b_0 + b_1 R_t + u_{2t}$$

حيث R معدل الفائدة ، M عرض النقود و Y الدخل (أ) لماذا يعتبر هذا نموذج معادلات آنية ؟ ماهى المتغيرات الداخلية والمتغيرات والمتغيرات الخارجية ؟ لماذا يعطى تقدير معادلات R و Y باستخدام OLS تقديرات معالم متحيزة وغير متسقة ؟

(ب) أوجد الشكل المختزل للنموذج . (جـ) هل هذا النموذج ناقص التمييز ، أو زائد التمييز ، أو مميز تماماً ؟ لماذا ؟ ماهى قيم المعاملات الهيكلية ؟ ماهو أسلوب التقدير المناسب للنموذج ؟ إشرح هذا الأسلوب . (د) لو أن المعادلة الأولى ، أو معادلة R تضمنت Y_{t-1} كتغير مفسر إضافى ، هل يعتبر هذا النموذج مميزاً ، زائد التمييز ، أو ناقص التمييز ؟ ماهى قيم معاملات الميل الهيكلية ؟ ماذا يعتبر أسلوباً مناسباً لتقدير النموذج ؟ إشرح هذا الأسلوب .

الإجابــات :

١ – (أ) أنظر جدول (٢)

$$\hat{b}_1 = \frac{\sum x_i y_i}{\sum x_i^2} = \frac{28}{34} \cong 0.82$$ (من الأعمدة السبعة الأولى فى جدول ٢)

$$\hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X} \cong 13 - (0.82)(8) \cong 6.44$$
$$\hat{Y}_i = 6.44 + 0.82 X_i$$

جدول (٢) مسودة

n	Y_i	X_i	y_i	x_i	$x_i y_i$	x_i^2	$\hat{Y}_i$	e_i	e_i^2	X_i^2	y_i^2
1	12	5	− 1	− 3	3	9	10.54	1.46	2.1316	25	1
2	14	11	1	3	3	9	15.46	− 1.46	2.1316	121	1
3	10	7	− 3	− 1	3	1	12.18	− 2.18	4.7524	49	9
4	13	8	0	0	0	0	13.00	0.00	0.0000	64	0
5	17	11	4	3	12	9	15.46	1.54	2.3716	121	16
6	12	7	− 1	− 1	1	1	12.18	− 0.18	0.0324	49	1
7	11	6	− 2	− 2	4	4	11.36	− 0.36	0.1296	36	4
8	15	9	2	1	2	1	13.82	1.18	1.3924	81	4
$n = 8$	$\sum Y_i = 104$ $\bar{Y} = 13$	$\sum X_i = 64$ $\bar{X} = 18$	$\sum y_i = 0$	$\sum x_i = 0$	$\sum x_i y_i = 28$	$\sum x_i = 34$		$\sum e_i = 0$	$\sum e_i^2 = 12.9416$	$\sum X_i^2 = 546$	$\sum y_i^2 = 36$

(ب) $$s_{\hat{b}_0}^2 = \frac{\sum e_i^2}{(n-k)} \frac{\sum X_i^2}{n \sum x_i^2} = \frac{(12.9416)(546)}{(8-2)(8)(34)} \cong 4.33$$ و $s_{\hat{b}_0} \cong 2.08$

$$s_{\hat{b}_1}^2=\frac{\sum e_i^2}{(n-k)\sum x_i^2}=\frac{12.9416}{(8-2)(34)}\cong 0.06 \quad \text{and} \quad s_{\hat{b}_1}\cong 0.25$$

$$t_0=\frac{\hat{b}_0}{s_{\hat{b}_0}}=\frac{6.44}{2.08}\cong 3.10$$ وهى معنوية عند مستوى 5%

$$t_1=\frac{\hat{b}_1}{s_{\hat{b}_1}}=\frac{0.82}{0.25}\cong 3.28$$ وهى أيضاً معنوية عند مستوى 5%

(جـ)

$$R^2=1-\frac{\sum e_i^2}{\sum y_i^2}=1-\frac{12.9416}{36}\cong 0.6405, \text{ or } 64.05\%$$

$$\hat{Y}_i=\underset{(3.10)}{6.44}+\underset{(3.28)}{0.82X_i} \qquad R^2\cong 64.05$$

(د)

$$\hat{Y}_F=6.44+0.82(10)=14.64$$

$$s_F^2=\frac{\sum e_i^2}{(n-2)}\left[1+\frac{1}{n}+\frac{(X_F-\bar{X})^2}{\sum x_i^2}\right]=\frac{12.9416}{6}\left[1+\frac{1}{8}+\frac{(10-8)^2}{34}\right]$$

$$s_F^2=2.67 \quad \text{and} \quad s_F\cong 1.63$$

وبالتالى ، فإن فـترة الثقة أو التنبؤ 95% بالنسبة إلى Y_F تكون $Y_F=14.64\pm 2.45\,(1.63)$ ، حيث $t_{0.052}=\pm 2.45$ عند درجات حرية $n-k=8-2=6\text{df}$ فنكون واثقين 95% أن $10.65\le Y_F\le 18.63$

٢ — (أ) من مسلمات نظرية طلب المستهلك أن الكمية المطلوبة من سلعة ما تتغير عكسياً مع سعر السلعة ولكن طردياً مع دخل المستهلك (إذا كانت السلعة عادية) ومع سعر السلع البديلة . وعليه فإن إشارات $\hat{b}_1$ و $\hat{b}_2$ تتفق مع ما تتوقعه نظرية الطلب ، بينما إشارة $\hat{b}_3$ تخالفها .

(ب)

$$t_1=-7/2=-3.5,\ t_2=2.4/0.8=3, \text{ and } t_3=4/18\cong 0.22.$$

وبالتالى فإن $\hat{b}_1$ و $\hat{b}_2$ معنوية إحصائياً عند مستوى 5% ، بينما $\hat{b}_3$ ليست كذلك .

(جـ)

$$R^2=\frac{\hat{b}_1\sum yx_1+\hat{b}_2\sum yx_2}{\sum y^2}=\frac{-7(10)+2.4(45)}{40}=\frac{-70+108}{40}=0.9500, \text{ or } 95\%$$

(د)

$$\bar{R}^2=1-(1-R^2)\frac{n-1}{n-4}=1-(1-0.95)\frac{23}{20}=1-(0.05)(1.15)=0.9425, \text{ or } 94.25\%$$

(هـ) حيث أن

$$F_{3,20} = \frac{R^2/k-1}{(1-R^2)/n-k} = \frac{0.95/4-1}{(1-0.95)/24-4} \cong \frac{0.3167}{0.0025} = 126.68$$

فإن R^2 تختلف معنوياً عن الصفر عند مستوى 5%

(و) حيث أن :

$$R^2 = 1-(\sum e^2/\sum y^2),\ \sum e^2 = (1-R^2)\sum y^2 = (1-0.95)(40) = 2.$$

فيكون

$$s = \sqrt{\frac{\sum e^2}{n-k}} = \sqrt{2/20} \cong 0.32$$

(ز)

$$\eta_{X_1} = \hat{b}_1(\bar{X}_1/\bar{Y}) = -7(8/32) = -1.75.\quad \eta_{X_2} = \hat{b}_2(\bar{X}_2/\bar{Y}) = 2.4(16/32) = 1.2$$

٣ – (أ) تعطى القيمة المنخفضة جداً لإحصائية ديربن – واتسون ، d ، الدليل على وجود الارتباط الذاتى . يشير الارتباط الذاتى إلى الحالة التى يكون فيها حد الخطأ فى فترة زمنية معينة مرتبطاً مع حد الخطأ فى أى فترة أخرى . الشكل الأكثر شيوعاً للارتباط الذاتى فى بيانات السلاسل الزمنية هو الارتباط الذاتى الموجب من الدرجة الأولى . فى وجود الارتباط الذاتى ، تظل معالم OLS غير متحيزة ومتسقة ، ولكن الأخطاء المعيارية لمعالم الانحدار المقدرة تكون متحيزة ، مما يؤدى إلى اختبارات إحصائية خاطئة وفترات ثقة متحيزة

(ب) يمكن الحصول على تقدير معامل الارتباط الذاتى ρ ، من معامل Y_{t-1} فى الانحدار التالى :

$$\hat{Y}_t = \hat{b}_0 + \hat{\rho}Y_{t-1} + \hat{b}_1X_{1t} - \hat{b}_1\rho X_{1t-1} + \hat{b}_2X_{2t} - \hat{b}_2\rho X_{t-1}$$

(جـ) يمكن إيجاد المتغيرات المحولة لتصحيح الارتباط الذاتى كالآتى (حيث تشير النجمة إلى المتغيرات المحولة) :

$$Y_t^* = Y_t - \hat{\rho}Y_{t-1},\quad X_{1t}^* = X_{1t} - \hat{\rho}X_{1t-1},\quad \text{and } X_{2t}^* = X_{2t} - \hat{\rho}X_{2t-1}$$

$$Y_1^* = Y_1\sqrt{1-\hat{\rho}^2}\quad X_{11}^* = X_1\sqrt{1-\hat{\rho}^2}\quad \text{and } X_{21} = X_2\sqrt{1-\hat{\rho}^2}$$

(د) حيث أن d تظل منخفضة جداً ، فإن الدليل على الارتباط الذاتى يبقى حتى بعد التعديل . فى هذه الحالة ، من المحتمل أن يكون الارتباط الذاتى راجعاً إلى أن بعض المتغيرات المفسرة الهامة لم تدخل فى الانحدار ، أو إلى شكل دالة غير ملائم ، أو بشكل أعم إلى تحديد متحيز للنموذج . وبالتالى ، فقبل تحويل المتغيرات فى محاولة للتغلب على الارتباط الذاتى ، من المهم أن يتضمن الانحدار كل المتغيرات ، وأن يستخدم شكل الدالة الذى تقترحه نظرية الاستثمار ، وبصفة عامة تجنب التحديد غير الصحيح للنموذج .

٤ – (أ) نموذج المعادلتين هذا آنى لأن R و Y تتحددان معاً . أى أن $R = f(Y)$ و $Y = f(R)$. والمتغيرات الداخلية فى النموذج هى R و Y . بينما M هى متغير خارجى أى يتحدد خارج النموذج . تقدير دالة R باستخدام OLS يعطى تقديرات متحيزة وغير متسقة لأن Y_t تكون مرتبطة مع x_{1t} . وبالمثل ، فإن تقدير المعادلة الثانية ، أى معادلة Y ، باستخدام OLS يعطى أيضاً تقديرات معالم متحيزة وغير متسقة لأن R و u_2 مرتبطان ببعضهما .

(ب) بالتعويض بقيمة Y المعطاة فى المعادلة الثانية فى المعادلة الأولى ، نحصل على :

$$R_t = a_0 + a_1 M_t + a_2(b_0 + b_1 R_t + u_{2t}) + u_{1t}$$

$$R_t - a_2 b_1 R_t = a_0 + a_2 b_0 + a_1 M_t + a_2 u_{2t} + u_{1t}$$

$$R_t = \frac{a_0 + a_2 b_0}{1 - a_2 b_1} + \frac{a_1}{1 - a_2 b_1} M_t + \frac{a_2 u_{2t} + u_{1t}}{1 - a_2 b_1} \quad \text{or} \quad R_t = \pi_0 + \pi_1 M_t + v_{1t}$$

بالتعويض بقيمة R_t المعطاة فى المعادلة الأولى فى المعادلة الثانية ، نحصل على :

$$Y_t = b_0 + b_1(a_0 + a_1 M_t + a_2 Y_t + u_{1t}) + u_{2t}$$

$$Y_t - a_2 b_1 Y_t = a_0 b_1 + b_0 + a_1 b_1 M_t + b_1 u_{1t} + u_{2t}$$

$$Y_t = \frac{a_0 b_1 + b_0}{1 - a_2 b_1} + \frac{a_1 b_1}{1 - a_2 b_1} M_t + \frac{b_1 u_{1t} + u_{2t}}{1 - a_2 b_1} \quad \text{or} \quad Y_t = \pi_2 + \pi_3 M_t + v_{2t}$$

(جـ) حيث أن المعادلة الأولى ، أى معادلة R ، لاتستبعد أى متغير خارجى ، فإنها ناقصة التمييز . وحيث أن عدد المتغيرات الخارجية المستبعدة من المعادلة الثانية ، أى معادلة Y (وعددها 1 ، أى المتغير M) يساوى عدد المتغيرات الداخلية (وهى R و Y) ناقصاً 1 ، فإن المعادلة الثانية ، أى معادلة Y ، مميزة بالضبط . $b_1 = \pi_3/\pi_1$ و $b_0 = \pi_2 - b_1\pi_0$. ولايمكن حساب a_1 و a_2 لأن معادلة R ناقصة التمييز . وتعتبر المربعات الصغرى غير المباشرة ILS أسلوباً مناسباً لتقدير المعادلة المميزة تماماً Y. ويتضمن هذا تقدير OLS لمعادلة الشكل المختزل R_t ثم استخدام $\hat{R}_t$ لتقدير المعادلة الهيكلية Y وعندما يتم ذلك تكون $\hat{b}_1$ متسقة .

(د) إذا تضمنت المعادلة الأولى ، أو معادلة R المتغير الإضافى Y_{t-1} ، تظل المعادلة الأولى ناقصة التمييز ، ولكن المعادلة الثانية تصبح زائدة التمييز . ويمكن تقدير قيمتين للمعلمة b_1 من معاملات الشكل المختزل ، ولكنه يكون من المستحيل حساب أى معامل ميل هيكلى لمعادلة R غير المميزة . وتكون المربعات الصغرى على مرحلتين (2SLS) أسلوباً ملائماً لتقدير معادلة Y زائدة التمييز . يتضمن هذا إيجاد انحدار R_t على M_t و Y_{t-1} أولا ، ثم استخدام R_t لتقدير معادلة Y الهيكلية . وعندما يتم هذا تكون $\hat{b}_1$ متسقة .

ملحق ١

احتمالات ذي الحدين

n	x	.01	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
1	0	.9900	.9500	.9000	.8500	.8000	.7500	.7000	.6500	.6000	.5500	.5000
	1	.0100	.0500	.1000	.1500	.2000	.2500	.3000	.3500	.4000	.4500	.5000
2	0	.9801	.9025	.8100	.7225	.6400	.5625	.4900	.4225	.3600	.3025	.2500
	1	.0198	.0950	.1800	.2550	.3200	.3750	.4200	.4550	.4800	.4950	.5000
	2	.0001	.0025	.0100	.0225	.0400	.0625	.0900	.1225	.1600	.2025	.2500
3	0	.9703	.8574	.7290	.6141	.5120	.4219	.3430	.2746	.2160	.1664	.1250
	1	.0294	.1354	.2430	.3251	.3840	.4219	.4410	.4436	.4320	.4084	.3750
	2	.0003	.0071	.0270	.0574	.0960	.1406	.1890	.2389	.2880	.3341	.3750
	3	.0000	.0001	.0010	.0034	.0080	.0156	.0270	.0429	.0640	.0911	.1250
4	0	.9606	.8145	.6561	.5220	.4096	.3164	.2401	.1785	.1296	.0915	.0625
	1	.0388	.1715	.2916	.3685	.4096	.4219	.4116	.3845	.3456	.2995	.2500
	2	.0006	.0135	.0486	.0975	.1536	.2109	.2646	.3105	.3456	.3675	.3750
	3	.0000	.0005	.0036	.0115	.0256	.0469	.0756	.1115	.1536	.2005	.2500
	4	.0000	.0000	.0001	.0005	.0016	.0039	.0081	.0150	.0256	.0410	.0625
5	0	.9510	.7738	.5905	.4437	.3277	.2373	.1681	.1160	.0778	.0503	.0312
	1	.0480	.2036	.3280	.3915	.4096	.3955	.3602	.3124	.2592	.2059	.1562
	2	.0010	.0214	.0729	.1382	.2048	.2637	.3087	.3364	.3456	.3369	.3125
	3	.0000	.0011	.0081	.0244	.0512	.0879	.1323	.1811	.2304	.2757	.3125
	4	.0000	.0000	.0004	.0022	.0064	.0146	.0284	.0488	.0768	.1128	.1562
	5	.0000	.0000	.0000	.0001	.0003	.0010	.0024	.0053	.0102	.0185	.0312
6	0	.9415	.7351	.5314	.3771	.2621	.1780	.1176	.0754	.0467	.0277	.0156
	1	.0571	.2321	.3543	.3993	.3932	.3560	.3025	.2437	.1866	.1359	.0938
	2	.0014	.0305	.0984	.1762	.2458	.2966	.3241	.3280	.3110	.2780	.2344
	3	.0000	.0021	.0146	.0415	.0819	.1318	.1852	.2355	.2765	.3032	.3125
	4	.0000	.0001	.0012	.0055	.0154	.0330	.0595	.0951	.1382	.1861	.2344
	5	.0000	.0000	.0001	.0004	.0015	.0044	.0102	.0205	.0369	.0609	.0938
	6	.0000	.0000	.0000	.0000	.0001	.0002	.0007	.0018	.0041	.0083	.0156
7	0	.9321	.6983	.4783	.3206	.2097	.1335	.0824	.0490	.0280	.0152	.0078
	1	.0659	.2573	.3720	.3960	.3670	.3115	.2471	.1848	.1306	.0872	.0547
	2	.0020	.0406	.1240	.2097	.2753	.3115	.3177	.2985	.2613	.2140	.1641
	3	.0000	.0036	.0230	.0617	.1147	.1730	.2269	.2679	.2903	.2918	.2734
	4	.0000	.0002	.0026	.0109	.0287	.0577	.0972	.1442	.1935	.2388	.2734
	5	.0000	.0000	.0002	.0012	.0043	.0115	.0250	.0466	.0774	.1172	.1641
	6	.0000	.0000	.0000	.0001	.0004	.0013	.0036	.0084	.0172	.0320	.0547
	7	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0006	.0016	.0037	.0078
8	0	.9227	.6634	.4305	.2725	.1678	.1002	.0576	.0319	.0168	.0084	.0039
	1	.0746	.2793	.3826	.3847	.3355	.2670	.1977	.1373	.0896	.0548	.0312
	2	.0026	.0515	.1488	.2376	.2936	.3115	.2065	.2587	.2090	.1569	.1094
	3	.0001	.0054	.0331	.0839	.1468	.2076	.2541	.2786	.2787	.2568	.2188
	4	.0000	.0004	.0046	.0185	.0459	.0865	.1361	.1875	.2322	.2627	.2734
	5	.0000	.0000	.0004	.0026	.0092	.0231	.0467	.0808	.1239	.1719	.2188
	6	.0000	.0000	.0000	.0002	.0011	.0038	.0100	.0217	.0413	.0403	.1094
	7	.0000	.0000	.0000	.0000	.0001	.0004	.0012	.0033	.0079	.0164	.0312
	8	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0007	.0017	.0039
9	0	.9135	.6302	.3874	.2316	.1342	.0751	.0404	.0207	.0101	.0046	.0020
	1	.0830	.2985	.3874	.3679	.3020	.2253	.1556	.1004	.0605	.0339	.0176
	2	.0034	.0629	.1722	.2597	.3020	.3003	.2668	.2162	.1612	.1110	.0703

n	x	.01	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
9	3	.0001	.0077	.0446	.1069	.1762	.2336	.2668	.2716	.2508	.2119	.1641
	4	.0000	.0006	.0074	.0283	.0661	.1168	.1715	.2194	.2508	.2600	.2461
	5	.0000	.0000	.0008	.0050	.0165	.0389	.0735	.1181	.1672	.2128	.2461
	6	.0000	.0000	.0001	.0006	.0028	.0087	.0210	.0424	.0743	.1160	.1641
	7	.0000	.0000	.0000	.0000	.0003	.0012	.0039	.0098	.0212	.0407	.0703
	8	.0000	.0000	.0000	.0000	.0000	.0001	.0004	.0013	.0035	.0083	.0176
	9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0008	.0020
10	0	.9044	.5987	.3487	.1969	.1074	.0563	.0282	.0135	.0060	.0025	.0010
	1	.0914	.3151	.3874	.3474	.2684	.1877	.1211	.0725	.0403	.0207	.0098
	2	.0042	.0746	.1937	.2759	.3020	.2816	.2335	.1757	.1209	.0763	.0439
	3	.0001	.0105	.0574	.1298	.2013	.2503	.2668	.2522	.2150	.1665	.1172
	4	.0000	.0010	.0112	.0401	.0881	.1460	.2001	.2377	.2508	.2384	.2051
	5	.0000	.0001	.0015	.0085	.0264	.0584	.1029	.1536	.2007	.2340	.2461
	6	.0000	.0000	.0001	.0012	.0055	.0162	.0368	.0689	.1115	.1596	.2051
	7	.0000	.0000	.0000	.0001	.0008	.0031	.0090	.0212	.0425	.0746	.1172
	8	.0000	.0000	.0000	.0000	.0001	.0004	.0014	.0043	.0106	.0229	.0439
	9	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0016	.0042	.0098
	10	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010
11	0	.8953	.5688	.3138	.1673	.0859	.0422	.0198	.0088	.0036	.0014	.0005
	1	.0995	.3293	.3835	.3248	.2362	.1549	.0932	.0518	.0266	.0125	.0054
	2	.0050	.0867	.2131	.2866	.2953	.2581	.1998	.1395	.0887	.0513	.0269
	3	.0002	.0137	.0710	.1517	.2215	.2581	.2568	.2254	.1774	.1259	.0806
	4	.0000	.0014	.0158	.0536	.1107	.1721	.2201	.2428	.2365	.2060	.1611
	5	.0000	.0001	.0025	.0132	.0388	.0803	.1321	.1830	.2207	.2360	.2256
	6	.0000	.0000	.0003	.0023	.0097	.0268	.0566	.0985	.1471	.1931	.2256
	7	.0000	.0000	.0000	.0003	.0017	.0064	.0173	.0379	.0701	.1128	.1611
	8	.0000	.0000	.0000	.0000	.0002	.0011	.0037	.0102	.0234	.0462	.0806
	9	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0018	.0052	.0126	.0269
	10	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0007	.0021	.0054
	11	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0005
12	0	.8864	.5404	.2824	.1422	.0687	.0317	.0138	.0057	.0022	.0008	.0002
	1	.1074	.3413	.3766	.3012	.2062	.1267	.0712	.0368	.0174	.0075	.0029
	2	.0060	.0988	.2301	.2924	.2835	.2323	.1678	.1088	.0639	.0339	.0161
	3	.0002	.0173	.0852	.1720	.2362	.2581	.2397	.1954	.1419	.0923	.0537
	4	.0000	.0021	.0213	.0683	.1329	.1936	.2311	.2367	.2128	.1700	.1208
	5	.0000	.0002	.0038	.0193	.0532	.1032	.1585	.2039	.2270	.2225	.1934
	6	.0000	.0000	.0005	.0040	.0155	.0401	.0792	.1281	.1766	.2124	.2256
	7	.0000	.0000	.0000	.0006	.0033	.0115	.0291	.0591	.1009	.1489	.1934
	8	.0000	.0000	.0000	.0001	.0005	.0024	.0078	.0199	.0420	.0762	.1208
	9	.0000	.0000	.0000	.0000	.0001	.0004	.0015	.0048	.0125	.0277	.0537
	10	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0008	.0025	.0068	.0161
	11	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010	.0029
	12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002
13	0	.8775	.5133	.2542	.1209	.0550	.0238	.0097	.0037	.0013	.0004	.0001
	1	.1152	.3512	.3672	.2774	.1787	.1029	.0540	.0259	.0113	.0045	.0016
	2	.0070	.1109	.2448	.2937	.2680	.2059	.1388	.0836	.0453	.0220	.0095
	3	.0003	.0214	.0997	.1900	.2457	.2517	.2181	.1651	.1107	.0660	.0349
	4	.0000	.0028	.0277	.0838	.1535	.2097	.2337	.2222	.1845	.1350	.0873

مثال : $P(\ X = 3, n = 5, P = 0.30\ ) = 0.1323$

(ملحق ١) احتمالات ذى الحدين

n	x	.01	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
								p				
13	5	.0000	.0003	.0055	.0266	.0691	.1258	.1803	.2154	.2214	.1989	.1571
	6	.0000	.0000	.0008	.0063	.0230	.0559	.1030	.1546	.1968	.2169	.2095
	7	.0000	.0000	.0001	.0011	.0058	.0186	.0442	.0833	.1312	.1775	.2095
	8	.0000	0000	.0001	.0001	.0011	.0047	.0142	.0336	.0656	.1089	.1571
	9	.0000	.0000	.0000	.0000	.0001	.0009	.0034	.0101	.0243	.0495	.0873
	10	.0000	.0000	.0000	.0000	.0000	.0001	.0006	.0022	.0065	.0162	0349
	11	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0012	.0036	.0095
	12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0016
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
14	0	.8687	.4877	.2288	.1028	.0440	.0178	.0068	.0024	.0008	.0002	.0001
	1	.1229	.3593	.3559	.2539	.1539	.0832	.0407	.0181	.0073	.0027	.0009
	2	.0081	.1229	.2570	.2912	.2501	.1802	.1134	.0634	.0317	.0141	.0056
	3	.0003	.0259	.1142	.2056	.2501	.2402	.1943	.1366	.0845	.0462	.0222
	4	.0000	.0037	.0349	.0998	.1720	.2202	.2290	.2022	.1549	.1040	.0611
	5	.0000	.0004	.0078	.0352	.0860	.1468	.1963	.2178	.2066	.1701	.1222
	6	.0000	.0000	.0013	.0093	.0322	.0734	.1262	.1759	.2066	.2088	.1833
	7	.0000	.0000	.0002	.0019	.0092	.0280	.0618	.1082	.1574	1952	.2095
	8	.0000	.0000	.0000	.0003	.0020	.0082	.0232	.0510	.0918	.1398	.1833
	9	.0000	.0000	.0000	.0000	.0003	.0018	.0066	.0183	.0408	.0762	.1222
	10	.0000	.0000	.0000	.0000	.0000	.0003	.0014	.0049	.0136	.0312	.0611
	11	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0010	.0033	.0093	.0222
	12	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0019	.0056
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0009
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
15	0	.8601	.4633	.2059	.0874	.0352	.0134	.0047	.0016	.0005	.0001	.0000
	1	.1303	.3658	.3432	.2312	.1319	.0668	.0305	.0126	.0047	.0016	.0005
	2	.0092	.1348	.2669	.2856	.2309	.1559	.0916	.0476	.0219	.0090	.0032
	3	.0004	.0307	.1285	.2184	.2501	.2252	.1700	.1110	.0634	.0318	.0139
	4	.0000	.0049	.0428	.1156	.1876	.2252	.2186	.1792	.1268	.0780	.0417
	5	.0000	.0006	.0105	.0449	.1032	.1651	.2061	.2123	.1859	.1404	.0916
	6	.0000	.0000	.0019	.0132	.0430	.0917	.1472	.1906	.2066	.1914	.1527
	7	.0000	.0000	.0003	.0030	.0138	.0393	.0811	.1319	.1771	.2013	.1964
	8	.0000	.0000	.0000	.0005	.0035	.0131	.0348	.0710	.1181	.1647	.1964
	9	.0000	.0000	.0000	.0001	.0007	.0034	.0116	.0298	.0612	.1048	.1527
	10	.0000	.0000	.0000	.0000	.0001	.0007	.0030	.0096	.0245	.0515	.0916
	11	.0000	.0000	.0000	.0000	.0000	.0001	.0006	.0024	.0074	.0191	.0417
	12	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0004	.0016	.0052	.0139
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010	.0032
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
16	0	.8515	.4401	.1853	.0743	.0281	.0100	.0033	.0010	.0003	.0001	.0000
	1	.1376	.3706	.3294	.2097	.1126	.0535	.0228	.0087	.0030	.0009	.0002
	2	.0104	.1463	.2745	.2775	.2111	.1336	.0732	.0353	.0150	.0056	.0018
	3	.0005	.0359	.1423	.2285	.2463	.2079	.1465	.0888	.0468	.0215	.0085
	4	.0000	.0061	.0514	.1311	.2001	.2262	.2040	.1553	.1014	.0572	.0278
	5	.0000	.0008	.0137	.0555	.1201	.1802	.2099	.2008	.1623	.1123	.0667
	6	.0000	.0001	.0028	.0180	.0550	.1101	.1649	.1982	.1983	.1684	.1222
	7	.0000	.0000	.0004	.0045	.0197	.0524	.1010	.1524	.1889	.1969	.1746
	8	.0000	.0000	.0001	.0009	.0055	.0197	.0487	.0923	.1417	.1812	.1964
	9	.0000	.0000	.0000	.0001	.0012	.0058	.0185	.0442	.0840	.1318	.1746
	10	.0000	.0000	.0000	.0000	.0002	.0014	.0056	.0167	.0392	.0755	.1222
	11	.0000	.0000	.0000	.0000	.0000	.0002	.0013	.0049	.0142	.0337	.0667
	12	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0011	.0040	.0115	.0278
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0008	.0029	.0085
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0018
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
17	0	.8429	.4181	.1668	.0631	.0225	.0075	.0023	.0007	.0002	.0000	.0000
	1	.1447	.3741	.3150	.1893	.0957	.0426	.0169	.0060	.0019	.0005	.0001
	2	.0117	.1575	.2800	.2673	.1914	.1136	.0581	.0260	.0102	.0035	.0010
	3	.0006	.0415	.1556	.2359	.2393	.1893	.1245	.0701	.0341	.0144	.0052
	4	.0000	.0076	.0605	.1457	.2093	.2209	.1868	.1320	.0796	.0411	.0182
	5	.0000	.0010	.0175	.0668	.1361	.1914	.2081	.1849	.1379	.0875	.0472
	6	.0000	.0001	.0039	.0236	.0680	.1276	.1784	.1991	.1839	.1432	.0944
	7	.0000	.0000	.0007	.0065	.0267	.0668	.1201	.1685	.1927	.1841	.1484
	8	.0000	.0000	.0001	.0014	.0084	.0279	.0644	.1134	.1606	.1883	.1855
	9	.0000	.0000	.0000	.0003	.0021	.0093	.0276	.0611	.1070	.1540	.1855
	10	.0000	.0000	.0000	.0000	.0004	.0025	.0095	.0263	.0571	.1008	.1484
	11	.0000	.0000	.0000	.0000	.0001	.0005	.0026	.0090	.0242	.0525	.0944
	12	.0000	.0000	.0000	.0000	.0000	.0001	.0006	.0024	.0081	.0215	.0472
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0021	.0068	.0182
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0004	.0016	.0052
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
18	0	.8345	.3972	.1501	.0536	.0180	.0056	.0016	.0004	.0001	.0000	.0000
	1	.1517	.3763	.3002	.1704	.0811	.0338	.0126	.0042	.0012	.0003	.0001
	2	.0130	.1683	.2835	.2556	.1723	.0958	.0458	.0190	.0069	.0022	.0006
	3	.0007	.0473	.1680	.2406	.2297	.1704	.1046	.0547	.0246	.0095	.0031
	4	.0000	.0093	.0700	.1592	.2153	.2130	.1681	.1104	.0614	.0291	.0117
	5	.0000	.0014	.0218	.0787	.1507	.1988	.2017	.1664	.1146	.0666	.0327
	6	.0000	.0002	.0052	.0301	.0816	.1436	.1873	.1941	.1655	.1181	.0708
	7	.0000	.0000	.0010	.0091	.0350	.0820	.1376	.1792	.1892	.1657	.1214
	8	.0000	.0000	.0002	.0022	.0120	.0376	.0811	.1327	.1734	.1864	.1669
	9	.0000	.0000	.0000	.0004	.0033	.0139	.0386	.0794	.1284	.1694	.1855
	10	.0000	.0000	.0000	.0001	.0008	.0042	.0149	.0385	.0771	.1248	.1669
	11	.0000	.0000	.0000	.0000	.0001	.0010	.0046	.0151	.0374	.0742	.1214
	12	.0000	.0000	.0000	.0000	.0000	.0002	.0012	.0047	.0145	.0354	.0708
	13	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0012	.0045	.0134	.0327
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0011	.0039	.0117
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0009	.0031
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0006
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
	18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

(ملحق ١)

احتمالات ذى الحدين

n	x	.01	.05	.10	.15	.20	.25	p .30	.35	.40	.45	.50
19	0	.8262	.3774	.1351	.0456	.0144	.0042	.0011	.0003	.0001	.0000	.0000
	1	.1586	.3774	.2852	.1529	.0685	.0268	.0093	.0029	.0008	.0002	.0000
	2	.0144	.1787	.2852	.2428	.1540	.0803	.0358	.0138	.0046	.0013	.0003
	3	.0008	.0533	.1796	.2428	.2182	.1517	.0869	.0422	.0175	.0062	.0018
	4	.0000	.0112	.0798	.1714	.2182	.2023	.1491	.0909	.0467	.0203	.0074
	5	.0000	.0018	.0266	.0907	.1636	.2023	.1916	.1468	.0933	.0497	.0222
	6	.0000	.0002	.0069	.0374	.0955	.1574	.1916	.1844	.1451	.0949	.0518
	7	.0000	.0000	.0014	.0122	.0443	.0974	.1525	.1844	.1797	.1443	.0961
	8	.0000	.0000	.0002	.0032	.0166	.0487	.0981	.1489	.1797	.1771	.1442
	9	.0000	.0000	.0000	.0007	.0051	.0198	.0514	.0980	.1464	.1771	.1762
	10	.0000	.0000	.0000	.0001	.0013	.0066	.0220	.0528	.0976	.1449	.1762
	11	.0000	.0000	.0000	.0000	.0003	.0018	.0077	.0233	.0532	.0970	.1442
	12	.0000	.0000	.0000	.0000	.0000	.0004	.0022	.0083	.0237	.0529	.0961
	13	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0024	.0085	.0233	.0518
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0006	.0024	.0082	.0222
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0022	.0074
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0018
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003
	18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	19	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
20	0	.8179	.3585	.1216	.0388	.0115	.0032	.0008	.0002	.0000	.0000	.0000
	1	.1652	.3774	.2702	.1368	.0576	.0211	.0068	.0020	.0005	.0001	.0000
	2	.0159	.1887	.2852	.2293	.1369	.0669	.0278	.0100	.0031	.0008	.0002
	3	.0010	.0596	.1901	.2428	.2054	.1339	.0716	.0323	.0123	.0040	.0011
	4	.0000	.0133	.0898	.1821	.2182	.1897	.1304	.0738	.0350	.0139	.0046
	5	.0000	.0022	.0319	.1028	.1746	.2023	.1789	.1272	.0746	.0365	.0148
	6	.0000	.0003	.0089	.0454	.1091	.1686	.1916	.1712	.1244	.0746	.0370
	7	.0000	.0000	.0020	.0160	.0545	.1124	.1643	.1844	.1659	.1221	.0739
	8	.0000	.0000	.0004	.0046	.0222	.0609	.1144	.1614	.1797	.1623	.1201
	9	.0000	.0000	.0001	.0011	.0074	.0271	.0654	.1158	.1597	.1771	.1602
	10	.0000	.0000	.0000	.0002	.0020	.0099	.0308	.0686	.1171	.1593	.1762
	11	.0000	.0000	.0000	.0000	.0005	.0030	.0120	.0336	.0710	.1185	.1602
	12	.0000	.0000	.0000	.0000	.0001	.0008	.0039	.0136	.0355	.0727	.1201
	13	.0000	.0000	.0000	.0000	.0000	.0002	.0010	.0045	.0146	.0366	.0739
	14	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0012	.0049	.0150	.0370
	15	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0003	.0013	.0049	.0148
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0003	.0013	.0046
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0011
	18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002
	19	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	20	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
25	0	.7778	.2774	.0718	.0172	.0038	.0008	.0001	.0000	.0000	.0000	.0000
	1	.1964	.3650	.1994	.0759	.0236	.0063	.0014	.0003	.0000	.0000	.0000
	2	.0238	.2305	.2659	.1607	.0708	.0251	.0074	.0018	.0004	.0001	.0000
	3	.0018	.0930	.2265	.2174	.1358	.0641	.0243	.0076	.0019	.0004	.0001
	4	.0001	.0269	.1384	.2110	.1867	.1175	.0572	.0224	.0071	.0018	.0004
25	5	.0000	.0060	.0646	.1564	.1960	.1645	.1030	.0506	.0199	.0063	.0016
	6	.0000	.0010	.0239	.0920	.1633	.1828	.1472	.0908	.0442	.0172	.0053
	7	.0000	.0001	.0072	.0441	.1108	.1654	.1712	.1327	.0800	.0381	.0143
	8	.0000	.0000	.0018	.0175	.0623	.1241	.1651	.1607	1200	.0701	.0322
	9	.0000	.0000	.0004	.0058	.0294	.0781	.1336	.1635	1511	.1084	.0609
	10	.0000	.0000	.0000	.0016	.0118	.0417	.0916	.1409	-.1612	.1419	.0974
	11	.0000	.0000	.0000	.0004	.0040	.0189	.0536	.1034	.1465	.1583	.1328
	12	.0000	.0000	.0000	.0000	.0012	.0074	.0268	.0650	.1140	.1511	.1550
	13	.0000	.0000	.0000	.0000	.0003	.0025	.0115	.0350	.0760	.1236	.1550
	14	.0000	.0000	.0000	.0000	.0000	.0007	.0042	.0161	.0434	.0867	.1328
	15	.0000	.0000	.0000	.0000	.0000	.0002	.0013	.0064	.0212	.0520	-.0974
	16	.0000	.0000	.0000	.0000	.0000	.0000	.0004	.0021	.0088	.0266	.0609
	17	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0006	.0031	.0115	.0322
	18	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0009	.0042	.0143
	19	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0013	.0053
	20	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0016
	21	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0004
	22	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
30	0	.7397	.2146	.0424	.0076	.0012	.0002	.0000	.0000	.0000	.0000	.0000
	1	.2242	.3389	.1413	.0404	.0093	.0018	.0003	.0000	.0000	.0000	.0000
	2	.0328	.2586	.2277	.1034	.0337	.0086	.0018	.0003	.0000	.0000	.0000
	3	.0031	.1270	.2361	.1703	.0785	.0269	.0072	.0015	.0003	.0000	.0000
	4	.0002	.0451	.1771	.2028	.1325	.0604	.0208	.0056	.0012	.0002	.0000
	5	.0000	.0124	.1023	.1861	.1723	.1047	.0464	.0157	.0041	.0008	.0001
	6	.0000	.0027	.0474	.1368	.1795	.1455	.0829	.0353	.0115	.0029	.0006
	7	.0000	.0005	.0180	.0828	.1538	.1662	.1219	.0652	.0263	.0081	.0019
	8	.0000	.0001	.0058	.0420	.1106	.1593	.1501	.1009	.0505	.0181	.0055
	9	.0000	.0000	.0016	.0181	.0676	.1298	.1573	.1328	.0823	.0382	.0133
	10	.0000	.0000	.0004	.0067	.0355	.0909	.1416	.1502	.1152	.0656	.0280
	11	.0000	.0000	.0001	.0022	.0161	.0551	.1103	.1471	.1396	.0976	.0509
	12	.0000	.0000	.0000	.0006	.0064	.0291	.0749	.1254	.1474	.1265	.0806
	13	.0000	.0000	.0000	.0001	.0022	.0134	.0444	.0835	.1360	.1433	.1115
	14	.0000	.0000	.0000	.0000	.0007	.0054	.0231	.0611	.1101	.1424	.1354
	15	.0000	.0000	.0000	.0000	.0002	.0019	.0106	.0351	.0783	.1242	.1445
	16	.0000	.0000	.0000	.0000	.0000	.0006	.0042	.0177	.0489	.0953	.1354
	17	.0000	.0000	.0000	.0000	.0000	.0002	.0015	.0079	.0269	.0642	.1115
	18	.0000	.0000	.0000	.0000	.0000	.0000	.0005	.0031	.0129	.0379	.0806
	19	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0010	.0054	.0196	.0509
	20	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0003	.0020	.0083	.0280
	21	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0006	.0034	.0133
	22	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	.0012	.0055
	23	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0003	.0019
	24	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0006
	25	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001

ملحق ٢

قيم $e^{-\lambda}$

λ	$e^{-\lambda}$	λ	$e^{-\lambda}$
0.0	1.00000	2.5	.08208
0.1	.90484	2.6	.07427
0.2	.81873	2.7	.06721
0.3	.74082	2.8	.06081
0.4	.67032	2.9	.05502
0.5	.60653	3.0	.04979
0.6	.54881	3.2	.04076
0.7	.49659	3.4	.03337
0.8	.44933	3.6	.02732
0.9	.40657	3.8	.02237
1.0	.36788	4.0	.01832
1.1	.33287	4.2	.01500
1.2	.30119	4.4	.01228
1.3	.27253	4.6	.01005
1.4	.24660	4.8	.00823
1.5	.22313	5.0	.00674
1.6	.20190	5.5	.00409
1.7	.18268	6.0	.00248
1.8	.16530	6.5	.00150
1.9	.14957	7.0	.00091
2.0	.13534	7.5	.00055
2.1	.12246	8.0	.00034
2.2	.00180	8.5	.00020
2.3	.10026	9.0	.00012
2.4	.09072	10.0	.00005

ملحق ٣

نسب مساحات
التوزيــع الطبيعى القيــاسى

المساحات المذكورة أدناه *

$z = \frac{X - \mu}{\sigma}$

μ z

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2518	.2549
0.7	.2580	2.612	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4014
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4983	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987									
3.5	.4997									
4.0	.4999									

مثال * : عند $z = 1.96$ ، المساحة المظللة 0.4750 من إجمالى المساحة 1.0000 .

ملحق ٤

جدول الأرقام العشوائية

10097	85017	84532	13618	23157	86952	02438	76520	91499	38631	79430	62421	97959	67422	69992	68479
37542	16719	82789	69041	05545	44109	05403	64894	80336	49172	16332	44670	35089	17691	89246	26940
08422	65842	27672	82186	14871	22115	86529	19645	44104	89232	57327	34679	62235	79655	81336	85157
99019	76875	20684	39187	38976	94324	43204	09376	12550	02844	15026	32439	58537	48274	81330	11100
12807	93640	39160	41453	97312	41548	93137	80157	63606	40387	65406	37920	08709	60623	2237	16505
66065	99478	70086	71265	11742	18226	29004	34072	61196	80240	44177	51171	08723	39323	05798	26457
31060	65119	26486	47353	43361	99436	42753	45571	15474	44910	99321	72173	56239	04595	10836	95270
85269	70322	21592	48233	93806	32584	21828	02051	94557	33663	86347	00926	44915	34823	51770	67897
63573	58133	41278	11697	49540	61777	67954	05325	42481	86430	19102	37420	41976	76559	24358	97344
73796	44655	81255	31133	36768	60452	38537	03529	23523	31379	68588	81675	15694	43438	36879	73208
98520	02295	13487	98662	07092	44673	61303	14905	04493	98086	32533	17767	14523	52494	24826	75246
11805	85035	54881	35587	43310	48897	48493	39808	00549	33185	04805	05431	94598	97654	16232	64051
83452	01197	86935	28021	61570	23350	65710	06288	35963	80951	68953	99634	81949	15307	00406	26898
88685	97907	19078	40646	31352	48625	44369	86507	59808	79752	02529	40200	73742	08391	49140	45427
99594	63268	96905	28797	57048	46359	74294	87517	46058	18633	99970	67348	49329	95236	32537	01390
65481	52841	59684	67411	09243	56092	84369	17468	32179	74029	74717	17674	90446	00597	45240	87379
80124	53722	71399	10916	07959	21225	13018	17727	69234	54178	10805	35635	45266	61406	41941	20117
74350	11434	51908	62171	93732	26958	02400	77402	19565	11664	77602	99817	28573	41430	96382	01758
69916	62375	99292	21177	72721	66995	07269	66252	45155	48324	32135	26803	16213	14938	71961	19476
09893	28337	20923	87929	61020	62841	31374	14225	94864	69074	45753	20505	78317	31984	98145	36168

ملحق ٥

المساحات المذكورة أدناه *

$$t = \frac{\overline{X} - \mu}{s_{\overline{X}}}$$

نسبة المساحة

(ذيل واحد)

$-\infty$ μ t $+\infty$

df	0.10	0.05	0.025	0.01	0.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898

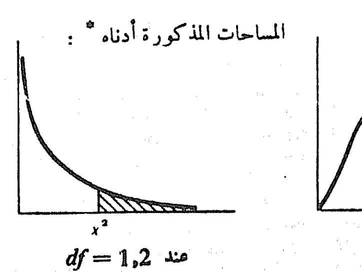
df	0.10	0.05	0.025	0.01	0.005
1[illegible]	[illegible]330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
40	1.303	1.684	2.021	2.423	2.704
60	1.296	1.671	2.000	2.390	2.660
120	1.289	1.658	1.980	2.358	2.617
∞	1.282	1.645	1.960	2.326	2.576

مثال * : بالنسبة المساحة المظللة والتي تمثل 0.05 من المساحة الكلية 1.10 قيمة t بدرجات حرية 10 هي 1.812

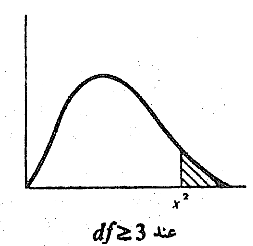
المصدر : من جدول ٣
Fisher and Yates, *Statisitical Tables for Biological, Agricultural and Medical Research*, 6th ed., 1974, Published by Longman Group Ltd., London (previously published by Oliver & Boyd, Edinburgh).

ملحق ٦

المساحات المذكورة أدناه * :



عند $df = 1,2$



عند $df \geq 3$

df	نسب مساحة توزيع كاى – تربيع										
	0.995	0.990	0.975	0.950	0.900	0.500	0.100	0.050	0.025	0.010	0.005
1	0.00004	0.00016	0.00098	0.00393	0.0158	0.455	2.71	3.84	5.02	6.63	7.88
2	0.0100	0.0201	0.0506	0.103	0.211	1.386	4.61	5.99	7.38	9.21	10.60
3	0.072	0.115	0.216	0.352	0.584	2.366	6.25	7.81	9.35	11.34	12.84
4	0.207	0.297	0.484	0.711	1.064	3.357	7.78	9.49	11.14	13.28	14.86
5	0.412	0.554	0.831	1.145	1.61	4.251	9.24	11.07	12.83	15.09	16.75
6	0.676	0.872	1.24	1.64	2.20	5.35	10.64	12.59	14.45	16.81	18.55
7	0.989	1.24	1.69	2.17	2.83	6.35	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	7.34	13.36	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	4.17	8.34	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	9.34	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	10.34	17.28	19.68	21.92	24.73	26.76
12	3.07	3.57	4.40	5.23	6.30	11.34	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	12.34	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	13.34	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.26	7.26	8.55	14.34	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	15.34	23.54	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	10.09	16.34	24.77	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	10.86	17.34	25.99	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	11.65	18.34	27.20	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	12.44	19.34	28.41	31.41	34.17	37.57	40.00
21	8.03	8.90	10.28	11.59	13.24	20.34	29.62	32.67	35.48	38.93	41.40
22	8.64	9.54	10.98	12.34	14.04	21.34	30.81	33.92	36.78	40.29	42.80
23	9.26	10.20	11.69	13.09	14.85	22.34	32.01	35.17	38.08	41.64	44.18
24	9.89	10.86	12.40	13.85	15.66	23.34	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	24.34	34.38	37.65	40.65	44.31	46.93
26	11.16	12.20	13.84	15.38	17.29	25.34	35.56	38.89	41.92	45.64	48.29
27	11.81	12.83	14.57	16.15	18.11	26.34	36.74	40.11	43.19	46.96	49.64
28	12.46	13.56	15.31	16.93	18.94	27.34	37.92	41.34	44.46	48.28	50.99
29	13.12	14.26	16.05	17.71	19.77	28.34	39.09	42.56	45.72	49.59	52.34
30	13.79	14.95	16.79	18.49	20.60	29.34	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	39.34	51.81	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	37.69	49.33	63.17	67.50	71.42	76.15	79.49
60	35.53	37.43	40.48	43.19	46.46	59.33	74.40	79.08	83.30	88.38	91.95
70	43.28	45.44	48.76	51.74	55.33	69.33	85.53	90.53	95.02	100.4	104.2
80	51.17	53.54	51.17	60.39	64.28	79.33	98.58	101.9	106.6	112.3	116.3
90	59.20	61.75	65.65	69.13	73.29	89.33	107.6	113.1	118.1	124.1	128.3
100	67.33	70.06	74.22	77.93	82.36	99.33	118.5	124.3	129.6	135.8	140.2

مثال * : بالنسبة للمساحة المظللة والتي تمثل 0.05 من المساحة الكلية 1.0 تحت دالة كثافة الاحتمال ، قيمة X^2 هي 18.31 عند درجات حرية df = 10

المصدر : من جدول ٤

Fisher and Yates, *Statistical Tables for Biological, Agriculturaland Medical Research*, 6th ed. 1974, published by Longman Group Ltd., London (previously published by Oliver & Boyd, Edinburgh).

ملحق ٧

قيم F باحتمالات 5% و 1%

المقام (columns) — البسط df (rows)

df	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞
1	161	200	216	225	230	234	237	239	241	242	243	244	245	246	248	249	250	251	252	253	253	254	254	254
	4,052	4,999	5,403	5,625	5,764	5,859	5,928	5,981	6,022	6,056	6,082	6,106	6,142	6,169	6,208	6,234	6,261	6,286	6,302	6,323	6,334	6,352	6,361	6,366
2	18.51	19.00	19.16	19.25	19.30	19.33	19.36	19.37	19.38	19.39	19.40	19.41	19.42	19.43	19.44	19.45	19.46	19.47	19.47	19.48	19.49	19.49	19.50	19.50
	98.49	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.41	99.42	99.43	99.44	99.45	99.46	99.47	99.48	99.48	99.49	99.49	99.49	99.50	99.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.88	8.84	8.81	8.78	8.76	8.74	8.71	8.69	8.66	8.64	8.62	8.60	8.58	8.57	8.56	8.54	8.54	8.53
	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.34	27.23	27.13	27.05	26.92	26.83	26.69	26.60	26.50	26.41	26.35	26.27	26.23	26.18	26.14	26.12
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.93	5.91	5.87	5.84	5.80	5.77	5.74	5.71	5.70	5.68	5.66	5.65	5.64	5.63
	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.54	14.45	14.37	14.24	14.15	14.02	13.93	13.83	13.74	13.69	13.61	13.57	13.52	13.48	13.46
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.78	4.74	4.70	4.68	4.64	4.60	4.56	4.53	4.50	4.46	4.44	4.42	4.40	4.38	4.37	4.36
	16.26	13.27	12.06	11.39	10.97	10.67	10.45	10.29	10.15	10.05	9.96	9.89	9.77	9.68	9.55	9.47	9.38	9.29	9.24	9.17	9.13	9.07	9.04	9.02
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.96	3.92	3.87	3.84	3.81	3.77	3.75	3.72	3.71	3.69	3.68	3.67
	13.74	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.79	7.72	7.60	7.52	7.39	7.31	7.23	7.14	7.09	7.02	6.99	6.94	6.90	6.88
7	5.59	4.74	4.34	4.12	3.97	3.87	3.79	3.73	3.68	3.63	3.60	3.57	3.52	3.49	3.44	3.41	3.38	3.34	3.32	3.29	3.28	3.25	3.24	3.23
	12.25	9.55	8.45	7.85	7.46	7.19	7.00	6.84	6.71	6.62	6.54	6.47	6.35	6.27	6.15	6.07	5.98	5.90	5.85	5.78	5.75	5.70	5.67	5.65
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.34	3.31	3.28	3.23	3.20	3.15	3.12	3.08	3.05	3.03	3.00	2.98	2.96	2.94	2.93
	11.26	8.65	7.59	7.01	6.63	6.37	6.19	6.03	5.91	5.82	5.74	5.67	5.56	5.48	5.36	5.28	5.20	5.11	5.06	5.00	4.96	4.91	4.88	4.86
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.13	3.10	3.07	3.02	2.98	2.93	2.90	2.86	2.82	2.80	2.77	2.76	2.73	2.72	2.71
	10.56	8.02	6.99	6.42	6.06	5.80	5.62	5.47	5.35	5.26	5.18	5.11	5.00	4.92	4.80	4.73	4.64	4.56	4.51	4.45	4.41	4.36	4.33	4.31
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.97	2.94	2.91	2.86	2.82	2.77	2.74	2.70	2.67	2.64	2.61	2.59	2.56	2.55	2.54
	10.04	7.56	6.55	5.99	5.64	5.39	5.21	5.06	4.95	4.85	4.78	4.71	4.60	4.52	4.41	4.33	4.25	4.17	4.12	4.05	4.01	3.96	3.93	3.91
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.86	2.82	2.79	2.74	2.70	2.65	2.61	2.57	2.53	2.50	2.47	2.45	2.42	2.41	2.40
	9.65	7.20	6.22	5.67	5.32	5.07	4.88	4.74	4.63	4.54	4.46	4.40	4.29	4.21	4.10	4.02	3.94	3.86	3.80	3.74	3.70	3.66	3.62	3.60
12	4.75	3.88	3.49	3.26	3.11	3.00	2.92	2.85	2.80	2.76	2.72	2.69	2.64	2.60	2.54	2.50	2.46	2.42	2.40	2.36	2.35	2.32	2.31	2.30
	9.33	6.93	5.95	5.41	5.06	4.82	4.65	4.50	4.39	4.30	4.22	4.16	4.05	3.98	3.86	3.78	3.70	3.61	3.56	3.49	3.46	3.41	3.38	3.36
13	4.67	3.80	3.41	3.18	3.02	2.92	2.84	2.77	2.72	2.67	2.63	2.60	2.55	2.51	2.46	2.42	2.38	2.34	2.32	2.28	2.26	2.24	2.22	2.21
	9.07	6.70	5.74	5.20	4.86	4.62	4.44	4.30	4.19	4.10	4.02	3.96	3.85	3.78	3.67	3.59	3.51	3.42	3.37	3.30	3.27	3.21	3.18	3.16
14	4.60	3.74	3.34	3.11	2.96	2.85	2.77	2.70	2.65	2.60	2.56	2.53	2.48	2.44	2.39	2.35	2.31	2.27	2.24	2.21	2.19	2.16	2.14	2.13
	8.86	6.51	5.56	5.03	4.69	4.46	4.28	4.14	4.03	3.94	3.86	3.80	3.70	3.62	3.51	3.43	3.34	3.26	3.21	3.14	3.11	3.06	3.02	3.00
15	4.54	3.68	3.29	3.06	2.90	2.79	2.70	2.64	2.59	2.55	2.51	2.48	2.43	2.39	2.33	2.29	2.25	2.21	2.18	2.15	2.12	2.10	2.08	2.07
	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.73	3.67	3.56	3.48	3.36	3.29	3.20	3.12	3.07	3.00	2.97	2.92	2.89	2.87
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.45	2.42	2.37	2.33	2.28	2.24	2.20	2.16	2.13	2.09	2.07	2.04	2.02	2.01
	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.61	3.55	3.45	3.37	3.25	3.18	3.10	3.01	2.96	2.98	2.86	2.80	2.77	2.75
17	4.45	3.59	3.20	2.96	2.81	2.70	2.62	2.55	2.50	2.45	2.41	2.38	2.33	2.29	2.23	2.19	2.15	2.11	2.08	2.04	2.02	1.99	1.97	1.96
	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.52	3.45	3.35	3.27	3.16	3.08	3.00	2.92	2.86	2.79	2.76	2.70	2.67	2.65
18	4.41	3.55	3.16	2.93	2.77	3.66	2.58	2.51	2.46	2.41	2.37	2.34	2.29	2.25	2.19	2.15	2.11	2.07	2.04	2.00	1.98	1.95	1.93	1.92
	8.28	6.01	5.09	4.58	4.25	4.01	3.85	3.71	3.60	3.51	3.44	3.37	3.27	3.19	3.07	3.00	2.91	2.83	2.78	2.71	2.68	2.62	2.59	2.57

(تابع)

(ملحق ٧)

قيم F

المقام df	البسط df: 1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞
19	4.38	3.52	3.13	2.90	2.74	2.63	2.55	2.48	2.43	2.38	2.34	2.31	2.26	2.21	2.15	2.11	2.07	2.02	2.00	1.96	1.94	1.91	1.90	1.88
	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.36	3.30	3.19	3.12	3.00	2.92	2.84	2.76	2.70	2.63	2.60	2.54	2.51	2.49
20	4.35	3.49	3.10	2.87	2.71	2.60	2.52	2.45	2.40	2.35	2.31	2.28	2.23	2.18	2.12	2.08	2.04	1.99	1.96	1.92	1.90	1.87	1.85	1.84
	8.10	5.85	4.94	4.43	4.10	3.87	3.71	3.56	3.45	3.37	3.30	3.23	3.13	3.05	2.94	2.86	2.77	2.69	2.63	2.56	2.53	2.47	2.44	2.42
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.28	2.25	2.20	2.15	2.09	2.05	2.00	1.96	1.93	1.89	1.87	1.84	1.82	1.81
	8.02	5.78	4.87	4.37	4.04	3.81	3.65	3.51	3.40	3.31	3.24	3.17	3.07	2.99	2.88	2.80	2.72	2.63	2.58	2.51	2.47	2.42	2.38	2.36
22	4.30	3.44	3.05	2.82	2.66	2.55	2.47	2.40	2.35	2.30	2.26	2.23	2.18	2.13	2.07	2.03	1.98	1.93	1.91	1.87	1.84	1.81	1.80	1.78
	7.94	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.18	3.12	3.02	2.94	2.83	2.75	2.67	2.58	2.53	2.46	2.42	2.37	2.33	2.31
23	4.28	3.42	3.03	2.80	2.64	2.53	2.45	2.38	2.32	2.28	2.24	2.20	2.14	2.10	2.04	2.00	1.96	1.91	1.88	1.84	1.82	1.79	1.77	1.76
	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.14	3.07	2.97	2.89	2.78	2.70	2.62	2.53	2.48	2.41	2.37	2.32	2.28	2.26
24	4.26	3.40	3.01	2.78	2.62	2.51	2.43	2.36	2.30	2.26	2.22	2.18	2.13	2.09	2.02	1.98	1.94	1.89	1.86	1.82	1.80	1.76	1.74	1.73
	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.25	3.17	3.09	3.03	2.93	2.85	2.74	2.66	2.58	2.49	2.44	2.36	2.33	2.27	2.23	2.21
25	4.24	3.38	2.99	2.76	2.60	2.49	2.41	2.34	2.28	2.24	2.20	2.16	2.11	2.06	2.00	1.96	1.92	1.87	1.84	1.80	1.77	1.74	1.72	1.71
	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.21	3.13	3.05	2.99	2.89	2.81	2.70	2.62	2.54	2.45	2.40	2.32	2.29	2.23	2.19	2.17
26	4.22	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15	2.10	2.05	1.99	1.95	1.90	1.85	1.82	1.78	1.76	1.72	1.70	1.69
	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.17	3.09	3.02	2.96	2.86	2.77	2.66	2.58	2.50	2.41	2.36	2.28	2.25	2.19	2.15	2.13
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.30	2.25	2.20	2.16	2.13	2.08	2.03	1.97	1.93	1.88	1.84	1.80	1.76	1.74	1.71	1.68	1.67
	7.68	5.49	4.60	4.11	3.79	3.56	3.39	3.26	3.14	3.06	2.98	2.93	2.83	2.74	2.63	2.55	2.47	2.38	2.33	2.25	2.21	2.16	2.12	2.10
28	4.20	3.34	2.95	2.71	2.56	2.44	2.36	2.29	2.24	2.19	2.15	2.12	2.06	2.02	1.96	1.91	1.87	1.81	1.78	1.75	1.72	1.69	1.67	1.65
	7.64	5.45	4.57	4.07	3.76	3.53	3.36	3.23	3.11	3.03	2.95	2.90	2.80	2.71	2.60	2.52	2.44	2.35	2.30	2.22	2.18	2.13	2.09	2.06
29	4.18	3.33	2.93	2.70	2.54	2.43	2.35	2.28	2.22	2.18	2.14	2.10	2.05	2.00	1.94	1.90	1.85	1.80	1.77	1.73	1.71	1.68	1.65	1.64
	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.08	3.00	2.92	2.87	2.77	2.68	2.57	2.49	2.41	2.32	2.27	2.19	2.15	2.10	2.06	2.03
30	4.17	3.32	2.92	2.69	2.53	2.42	2.34	2.27	2.21	2.16	2.12	2.09	2.04	1.99	1.93	1.89	1.84	1.79	1.76	1.72	1.69	1.66	1.64	1.62
	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.06	2.98	2.90	2.84	2.74	2.66	2.55	2.47	2.38	2.29	2.24	2.16	2.13	2.07	2.03	2.01
32	4.15	3.30	2.90	2.67	2.51	2.40	2.32	2.25	2.19	2.14	2.10	2.07	2.02	1.97	1.91	1.86	1.82	1.76	1.74	1.69	1.67	1.64	1.61	1.59
	7.50	5.34	4.46	3.97	3.66	3.42	3.25	3.12	3.01	2.94	2.86	2.80	2.70	2.62	2.51	2.42	2.34	2.25	2.20	2.12	2.08	2.02	1.98	1.96
34	4.13	3.28	2.88	2.65	2.49	2.38	2.30	2.23	2.17	2.12	2.08	2.05	2.00	1.95	1.89	1.84	1.80	1.74	1.71	1.67	1.64	1.61	1.59	1.57
	7.44	5.29	4.42	3.93	3.61	3.38	3.21	3.08	2.97	2.89	2.82	2.76	2.66	2.58	2.47	2.38	2.30	2.21	2.15	2.08	2.04	1.98	1.94	1.91
36	4.11	3.26	2.86	2.63	2.48	2.36	2.28	2.21	2.15	2.10	2.06	2.03	1.98	1.93	1.87	1.82	1.78	1.72	1.69	1.65	1.62	1.59	1.56	1.55
	7.39	5.25	4.38	3.89	3.58	3.35	3.18	3.04	2.94	2.86	2.78	2.72	2.62	2.54	2.43	2.35	2.26	2.17	2.12	2.04	2.00	1.94	1.90	1.87
38	4.10	3.25	2.85	2.62	2.46	2.35	2.26	2.19	2.14	2.09	2.05	2.02	1.96	1.92	1.85	1.80	1.76	1.71	1.67	1.63	1.60	1.57	1.54	1.53
	7.35	5.21	4.34	3.86	3.54	3.32	3.15	3.02	2.91	2.82	2.75	2.69	2.59	2.51	2.40	2.32	2.22	2.14	2.08	2.00	1.97	1.90	1.86	1.84
40	4.07	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.07	2.04	2.00	1.95	1.90	1.84	1.79	1.74	1.69	1.66	1.61	1.59	1.55	1.53	1.51
	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.88	2.80	2.73	2.66	2.56	2.49	2.37	2.29	2.20	2.11	2.05	1.97	1.94	1.88	1.84	1.81

(تـابع)

(ملحق ٧)

قيم F

المقام df	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞
42	4.07	3.22	2.83	2.59	2.44	2.32	2.24	2.17	2.11	2.06	2.02	1.99	1.94	1.89	1.82	1.78	1.73	1.68	1.64	1.60	1.57	1.54	1.51	1.49
	7.27	5.15	4.29	3.80	3.49	3.26	3.10	2.96	2.86	2.77	2.70	2.64	2.54	2.46	2.35	2.26	2.17	2.08	2.02	1.94	1.91	1.85	1.80	1.78
44	4.06	3.21	2.82	2.58	2.43	2.31	2.23	2.16	2.10	2.05	2.01	1.9[illegible]	1.92	1.88	1.81	1.76	1.72	1.66	1.63	1.58	1.56	1.52	1.50	1.48
	7.24	5.12	4.26	3.78	3.46	3.24	3.07	2.94	2.84	2.75	2.68	2.62	2.52	2.44	2.32	2.24	2.15	2.06	2.00	1.92	1.88	1.82	1.78	1.75
46	4.05	3.20	2.81	2.57	2.42	2.30	2.22	2.14	2.09	2.04	2.00	1.97	1.91	1.87	1.80	1.75	1.71	1.65	1.62	1.57	1.54	1.51	1.48	1.46
	7.21	5.10	4.24	3.76	3.44	3.22	3.05	2.92	2.82	2.73	2.66	2.60	2.50	2.42	2.30	2.22	2.13	2.04	1.98	1.90	1.86	1.80	1.76	1.72
48	4.04	3.19	2.80	2.56	2.41	2.30	2.21	2.14	2.08	2.03	1.99	1.96	1.90	1.86	1.79	1.74	1.70	1.64	1.61	1.56	1.53	1.50	1.47	1.45
	7.19	5.08	4.22	3.74	3.42	3.20	3.04	2.90	2.80	2.71	2.64	2.58	2.48	2.40	2.28	2.20	2.11	2.02	1.96	1.88	1.84	1.78	1.73	1.70
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.02	1.98	1.95	1.90	1.85	1.78	1.74	1.69	1.63	1.60	1.55	1.52	1.48	1.46	1.44
	7.17	5.06	4.20	3.72	3.41	3.18	3.02	2.88	2.78	2.70	2.62	2.56	2.46	2.39	2.26	2.18	2.10	2.00	1.94	1.86	1.82	1.76	1.71	1.68
55	4.02	3.17	2.78	2.54	2.38	2.27	2.18	2.11	2.05	2.00	1.97	1.93	1.88	1.83	1.76	1.72	1.67	1.61	1.58	1.52	1.50	1.46	1.43	1.41
	7.12	5.01	4.16	3.68	3.37	3.15	2.98	2.85	2.75	2.66	2.59	2.53	2.43	2.35	2.23	2.15	2.06	1.96	1.90	1.82	1.78	1.71	1.66	1.64
60	4.00	3.15	2.76	2.52	2.37	2.25	2.17	2.10	2.04	1.99	1.95	1.92	1.86	1.81	1.[illegible]	1.70	1.65	1.59	1.56	1.50	1.48	1.44	1.41	1.39
	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.56	2.50	2.40	2.32	2.2[illegible]	2.12	2.03	1.93	1.87	1.79	1.74	1.68	1.63	1.60
65	3.99	3.14	2.75	2.51	2.36	2.24	2.15	2.08	2.02	1.98	1.94	1.90	1.85	1.80	1.73	1.68	1.63	1.57	1.54	1.49	1.46	1.42	1.39	1.37
	7.04	4.95	4.10	3.62	3.31	3.09	2.93	2.79	2.70	2.61	2.54	2.47	2.37	2.30	2.18	2.09	2.00	1.90	1.84	1.76	1.71	1.64	1.60	1.56
70	3.98	3.13	2.74	2.50	2.35	2.23	2.14	2.07	2.01	1.97	1.93	1.89	1.84	1.79	1.72	1.67	1.62	1.56	1.53	1.47	1.45	1.40	1.37	1.35
	7.01	4.92	4.08	3.60	3.29	3.07	2.91	2.77	2.67	2.59	2.51	2.45	2.35	2.28	2.15	2.07	1.98	1.88	1.82	1.74	1.69	1.62	1.56	1.53
80	3.96	3.11	2.72	2.48	2.33	2.21	2.12	2.05	1.99	1.95	1.91	1.88	1.82	1.77	1.70	1.65	1.60	1.54	1.51	1.45	1.42	1.38	1.35	1.32
	6.96	4.88	4.04	3.56	3.25	3.04	2.87	2.74	2.64	2.55	2.48	2.41	2.32	2.24	2.11	2.03	1.94	1.84	1.78	1.70	1.65	1.57	1.52	1.49
100	3.94	3.09	2.70	2.46	2.30	2.19	2.10	2.03	1.97	1.92	1.88	1.85	1.79	1.75	1.68	1.63	1.57	1.51	1.48	1.42	1.39	1.34	1.30	1.28
	6.90	4.82	3.98	3.51	3.20	2.99	2.82	2.69	2.59	2.51	2.43	2.36	2.26	2.19	2.06	1.98	1.89	1.79	1.73	1.64	1.59	1.51	1.46	1.43
125	3.92	3.07	2.68	2.44	2.29	2.17	2.08	2.01	1.95	1.90	1.86	1.83	1.77	1.72	1.65	1.60	1.55	1.49	1.45	1.39	1.36	1.31	1.27	1.25
	6.84	4.78	3.94	3.47	3.17	2.95	2.79	2.65	2.56	2.47	2.40	2.33	2.23	2.15	2.03	1.94	1.85	1.75	1.68	1.59	1.54	1.46	1.40	1.37
150	3.91	3.06	2.67	2.43	2.27	2.16	2.07	2.00	1.94	1.89	1.85	1.82	1.76	1.71	1.64	1.59	1.54	1.47	1.44	1.37	1.34	1.29	1.25	1.22
	6.81	4.75	3.91	3.44	3.14	2.92	2.76	2.62	2.53	2.44	2.37	2.30	2.20	2.12	2.00	1.91	1.83	1.72	1.66	1.56	1.51	1.43	1.37	1.33
200	3.89	3.04	2.65	2.41	2.26	2.14	2.05	1.98	1.92	1.87	1.83	1.80	1.74	1.69	1.62	1.57	1.52	1.45	1.42	1.35	1.32	1.26	1.22	1.19
	6.76	4.71	3.88	3.41	3.11	2.90	2.73	2.60	2.50	2.41	2.34	2.28	2.17	2.09	1.97	1.88	1.79	1.69	1.62	1.53	1.48	1.39	1.33	1.28
400	3.86	3.02	2.62	2.39	2.23	2.12	2.03	1.96	1.90	1.85	1.91	1.78	1.72	1.67	1.60	1.54	1.49	1.42	1.38	1.32	1.28	1.22	1.16	1.13
	6.70	4.66	3.83	3.36	3.06	2.85	2.69	2.55	2.46	2.37	2.29	2.23	2.12	2.04	1.92	1.84	1.74	1.64	1.57	1.47	1.42	1.32	1.24	1.19
1000	3.85	3.00	2.61	2.38	2.22	2.10	2.02	1.95	1.89	1.84	1.80	1.76	1.70	1.65	1.58	1.53	1.47	1.41	1.36	1.30	1.26	1.19	1.13	1.08
	6.66	4.62	3.80	3.34	3.04	2.82	2.66	2.53	2.43	2.34	2.26	2.20	2.09	2.01	1.89	1.81	1.71	1.61	1.54	1.44	1.38	1.28	1.19	1.11
∞	3.84	2.99	2.60	2.37	2.21	2.09	2.01	1.94	1.88	1.83	1.79	1.75	1.69	1.64	1.57	1.52	1.46	1.40	1.35	1.28	1.24	1.17	1.11	1.00
	6.64	4.60	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.24	2.18	2.07	1.99	1.87	1.79	1.69	1.59	1.52	1.41	1.36	1.25	1.15	1.00

(Column heading: البسط df)

المصدر : طبع بتصريح من *Statistical Methods*. 6th ed., by George W. Snedecor and William G. Cochran. (c) 1967, by the Iowa State University Press, Ames, Iowa.

ملحق ٨

إحصائية ديربين واتسون

Significance Points of d_L and d_U: 5%

n	$k'=1$ d_L	$k'=1$ d_U	$k'=2$ d_L	$k'=2$ d_U	$k'=3$ d_L	$k'=3$ d_U	$k'=4$ d_L	$k'=4$ d_U	$k'=5$ d_L	$k'=5$ d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

Significance Points of d_L and d_U: 1%

n	$k'=1$ d_L	$k'=1$ d_U	$k'=2$ d_L	$k'=2$ d_U	$k'=3$ d_L	$k'=3$ d_U	$k'=4$ d_L	$k'=4$ d_U	$k'=5$ d_L	$k'=5$ d_U
15	0.81	1.07	0.70	1.25	0.59	1.46	0.49	1.70	0.39	1.96
16	0.84	1.09	0.74	1.25	0.63	1.44	0.53	1.66	0.44	1.90
17	0.87	1.10	0.77	1.25	0.67	1.43	0.57	1.63	0.48	1.85
18	0.90	1.12	0.80	1.26	0.71	1.42	0.61	1.60	0.52	1.80
19	0.93	1.13	0.83	1.26	0.74	1.41	0.65	1.58	0.56	1.77
20	0.95	1.15	0.86	1.27	0.77	1.41	0.68	1.57	0.60	1.74
21	0.97	1.16	0.89	1.27	0.80	1.41	0.72	1.55	0.63	1.71
22	1.00	1.17	0.91	1.28	0.83	1.40	0.75	1.54	0.66	1.69
23	1.02	1.19	0.94	1.29	0.86	1.40	0.77	1.53	0.70	1.67
24	1.04	1.20	0.96	1.30	0.88	1.41	0.80	1.53	0.72	1.66
25	1.05	1.21	0.98	1.30	0.90	1.41	0.83	1.52	0.75	1.65
26	1.07	1.22	1.00	1.31	0.93	1.41	0.85	1.52	0.78	1.64
27	1.09	1.23	1.02	1.32	0.95	1.41	0.88	1.51	0.81	1.63
28	1.10	1.24	1.04	1.32	0.97	1.41	0.90	1.51	0.83	1.62
29	1.12	1.25	1.05	1.33	0.99	1.42	0.92	1.51	0.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	0.94	1.51	0.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	0.96	1.51	0.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	0.98	1.51	0.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	0.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	0.95	1.59
35	1.19	1.31	1.14	1.37	1.08	1.44	1.03	1.51	0.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	0.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.20	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.61
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

ملحوظة : k' = عدد المتغيرات المفسرة باستثناء الحد الثابت

المصدر : J. Durbin and G.S. Waston "Testing for Serial Correlation in Least Squarez Regression" *Bionmetrika,* vol. 38, 1951 pp. 159-77

المصطلحات العلمية (عربى — انجليزى)

(أ)

اتساق	Consistency
احتمال	Probability
الاحتمال التجريبى	Empirical probability
احتمال شخصى	Personalistic (subjective) probability
الاحتمال الشرطى	Conditional probability
احتمال عدم الحدوث	Nonoccurance probability
الاحتمال المسبق (الكلاسيكى أو النظرى)	A priori (classical) probability
احتمال مشترك	Joint probability
أحداث متنافية	Mutually exclusive events
إحصاء استدلالى	Inferential statistics
إحصاء وصفى	Descriptive statistics
إحصائية	Statistic
اختبارات الفروض	Hypothesis testing
اختلاف التباين	Heteroscedasticity
الارتباط الذاتى	Autocorrelation (serial correlation)
ارتباط الرتب	Rank correlation
ارتباط متسلسل (ارتباط ذاتى)	Serial correlation
أفضل مقدرات خطية غير متحيزة	BLUE (Best Linear Unbiased Estimators)
اقتصاد قياسى	Econometrics
التـــواء	Skewness
الانحراف المعيارى	Standard deviation

(ب)

البواقى	Residuals
البواقى غير المفسرة	Unexplained reasiduals
بيانات غير مبوبة	Ungrouped data
بيانات مبوبة	Grouped data
بيانات مقطعية (كبيانات ميزانية الأسرة)	Cross-sectional data

(ت)

التبـــادل	Permurations
تبـــاين	Variance
تحليل الانحدار	Regression analysis
تحليل التباين	Analysis of variance
تحليل السلاسل الزمنية	Time-series analysis
تحـــيز	Bias
الترميز	Coding
تشتت	Dispersion
التشتت النسبي	Relative dispersion
تشويش	Disturbance
تعدد العلاقات الخطية	Multicollinearity
التغير المفسر ، المشروح	Explained variation
تفرطح	Kurtosis
تقـــدير	Estimate, Estimator
التقدير بنقطة	Point estimate
تقدير غير متحيز	Unbiased estimate
تكرار متجمع	Cumulative frequency
التكرار النسبي	Relative frequency
التمـــاثل	Symmetry
التنبـــؤ	Forecasting
التوافيق	Combinations
التوزيع الاحتمالى	Probability distribution
توزيع أسى	Exponential distribution
توزيع بواسون	Poisson distribution
توزيع تكرارى	Frequency distribution
توزيع ذى الحدين	Binomial distribution
توزيع طبيعى	Normal distribution
توزيع منفصل	Discrete distribution

(ج)

جدول الاقتران	Contingency table
جودة التوفيق	Goodness of fit

(ح)

حد الخطأ	Error term

(خ)

خطأ من النوع الأول	Type I Error
خطأ من النوع الثانى	Type II Error

(د)

دالة كثافة الاحتمال	Density function
دالة كثيرة الحدود	Polynomial function
دالة نصف لوغاريتمية	Semilog function
درجات الحرية	Degrees of freedom

(ر)

ربيع	Quartile

(ش)

شكل الانتشار	Scatter diagram
الشكل المختزل	Reduced form

(ع)

عدم تحيز فى المدى البعيد	Asymptotic unbiasedness
عشوائى (احتمالى)	Stocnastic
عشيرات	Deciles
علم الإحصاء	Statistics
عينة	Sample

(غ)

غير خطى	Nonlinear

(ف)

فترة ثقـــة	Confidence interval
الفرض البديل	Alternative hypothesis
الفرض العدمى	Null hypothesis
فضاء العينة	Sample space

(ق)

القيمة المتوقعة	Expected value

(م)

Variable	متغير
Dependent variable	متغير تابع
Endogenous variable	متغير داخلى
Exogenous variable	متغير خارجى
Dummy variable	متغير صورى
Random variable	متغير عشوائى
Qualitative variable	متغير كيفى
Lagged variable	متغير مبطأ
Independent variable	متغير مستقل
Explanatory variable	متغير مفسر
Instrumental variable	متغير وسيط
Weighted mean	متغير مرجح
Population	المجتمع
Infinite population	مجتمع غير محدود
Finite population	مجتمع محدود
Leptokurtic	مدبب
Histogram	المدرج التكرارى
Range	المـــدى
Interquartile range	المدى الربيعى
Ordinary Least Squares (OLS)	المربعات الصغرى العادية
Indirect Least Squares (ILS)	المربعات الصغرى غير المباشرة
Price elasticity	المرونة السعرية
Confidence level	مستوى ثقة
Frequency polygon	مضلع تكرارى
Simultaneous-equations	معادلات آنيـــة
Behavioral (structural) equations	معادلات سلوكية (هيكلية)
Variation coefficient	معامل الاختلاف
Correlation coefficient	معامل الارتباط
Partial correlation coefficient	معامل الارتباط الجزئى
Determination coefficient	معامل التحديد
Adjusted R^2	معامل التحديد المعدل
Coefficients	المعاملات
Structural coefficients	معاملات هيكلية
Sampling	معاينة
Stratified sampling	معاينة طبقية
Random sampling	معاينة عشوائية

Cluster sampling	معاينة عنقودية
Systematic sampling	معاينة منتظمة
Peramter	معلمة (وجمعها معالم)
Platykurtic	مفرطح
Estimater	مقـــدر
Operating Characteristic (OC) curve	منحنى توصيف العمليات
Ogive	المنحنى المتجمع
Critical region	المنطقة الحرجة
Rejection region	منطقة الرفض
Acceptance region	منطقة القبول
Percentiles	المئينـــات

(ن)

Central tendency	النزعة المركزية
Quartile deviation	نصف المدى الربيعى
Chebyshev's theorem (inequality)	نظرية (متباينة) تشيبشيف
Set theory	نظرية المجموعات
Central limit theorem	نظرية النهاية المركزية
Distributed lag model	نموذج إبطاء موزع
Harmonic mean	الوسط التوافقى
Geometric mean	الوسط الهندسى

المصطلحات العلمية (انجليزى — عربى)

(A)

A priori (classical) probability	الاحتمال المسبق (الكلاسيكى) ، النظرى
Acceptance region	منطقة القبول
Adjusted R^2	معامل التحديد المعدل
Alternative hyptohesis	الفرض البديل
Analysis of variance	تحليل التباين
Asymptotic unbiasedness	عدم تحيز فى المدى البعيد
Autocorrelation (serial correlation)	الارتباط الذاتى

(B)

Behavioral (structural) equations	معادلات سلوكية (هيكلية)
Bias	تحـيز
Binomial distribution	توزيع ذى الحدين
BLUE (best Linear Unbiased Estimators)	أفضل مقدرات خطية غير متحيزة

(C)

Central limit theorem	نظرية النهاية المركزية
Central tendency	النزعة المركزية
Chebyshev's theorem (inequality)	نظرية (متباينة) تشيبشيف
Cluster sampling	معاينة عنقودية
Coding	ترميز
Coefficients	معاملات
Combinations	توافيق
Conditional probability	احتمال شرطى
Confidence interval	فـترة ثقـة
Confidence level	مستوى ثقـة
Consistency	اتساق
Contingency table	جدول اقتران
Correlation coefficient	معامل ارتباط
Critical region	منطقة حرجة
Cross-sectional data	بيانات مقطعية (كبيانات ميزانية الأسرة)
Cumulative frequency	تكرار متجمع

(D)

Deciles	عشيرات
Degrees of freedom	درجات الحرية
Density function	دالة كثافة الاحتمال
Dependent variable	متغير تابع
Descriptive statistics	إحصاء وصفى
Determination coefficient	معامل التحديد
Discrete distribution	توزيع منفصل
Dispersion	تشتت
Distributed lag models	نماذج إبطاء موزعة
Disturbance	تشويش ، خطأ
Dummy variable	متغير صورى

(E)

Econometrics	اقتصاد قياسى
Empirical probability	الاحتمال التجريبى
Endogeneous variable	متغير داخلى
Error term	حد الخطأ
Estimate, Estimation	تقدير
Estimator	مقدر
Exactly identified equation	معادلة مميزة بالضبط
Exogeneous variable	متغير خارجى
Expected value	القيمة المتوقعة
Explained variation	التغير المفسر ، المشروح
Explanatory variable	متغير مفسر
Exponential distribution	توزيع أسى

(F)

Finitie population	مجتمع محدود
First-order autocorrelation	ارتباط ذاتى من الدرجة الأولى
Forecasting	التنبؤ
Frequency distribution	توزيع تكرارى
Frequency polygon	مضلع تكرارى

(G)

Geometric mean	الوسط الهندسى
Goodness of fit	جودة التوفيق
Gouped data	بيانات مبوبة

(H)

English	العربية
Harmonic mean	الوسط التوافقى
Heteroscedasticity	اختلاف التباين
Histogram	المدرج التكرارى
Hypothesis testing	اختبارات الفروض

(I)

English	العربية
Identification	تمييز
Independent variable	متغير مستقل
Indirect Least Squares (ILS)	المربعات الصغرى غير المباشرة
Inferential statistics	إحصاء استدلالى
Infinite population	مجتمع غير محدود
Instrumental variable	متغير وسيط
Interquartile range	المدى الربيعى

(J)

English	العربية
Joint probability	احتمال مشترك

(K)

English	العربية
Kurtosis	تفرطح

(L)

English	العربية
Lagged variable	متغير مبطأ
Leptokurtic	مدبب

(M)

English	العربية
Multicolinearity	تعدد علاقات خطية
Mutually exclusive events	أحداث متنافية

(N)

English	العربية
Nonlinear	غير خطى
Nonoccurance probability	احتمال عدم الحدوث
Normal distribution	توزيع طبيعى
Null hypothesis	فرض عدمى

(O)

English	العربية
Ogive	المنحنى المتجمع
Operating Characteristic (OC)curve	منحنى توصيف العمليات
Order condition	شرط الترتيب
Ordinary Least Squares (OLS)	المربعات الصغرى العادية
Overidentified equation	معادلة زائدة التمييز

(P)

Parameter	معلمـــة
Partial correlation coefficient	معامل الارتباط الجزئى
Percentiles	المئينـــات
Permutations	التباديل
Personalistic (subjective) probability	احتمال شخصى (ذاتى)
Platykurtic	مفرطح
Point estimate	التقدير بنقطة
Poisson distribution	توزيع بواسون
Polynomial function	دالة كثيرة الحدود
Population	المجتمع
Price elasticity	المرونة السعرية
Probability	الاحـــتمال
Probability distribution	التوزيع الاحتمالى

(Q)

Qualitative variable	متغير كيفى
Quartile	ربيـــع
Quartile deviation	نصف المدى الربيعى

(R)

Random Sampling	معاينة عشوائية
Random variable	متغير عشوائى
Range	مدى
Rank condition	شرط الرتبـــة
Rank correlation	ارتباط الرتب
Recursive model	نموذج متواتر
Reduced form	الشكل المختزل
Regression analysis	تحليل الانحـــدار
Rejection region	منطقة الرفض
Relative dispersion	تشتت نسبى
Relative frequency	تكرار نسبى
Residuals	بواقى

(S)

Sample	عينـــة
Sample space	فضاء العينـــة

Sampling	معـــاينة
Scatter diagram	شكل الانتشار
Semilog function	دالة نصف لوغاريتمية
Serial correlation	ارتباط متسلسل
Set theory	نظرية المجموعات
Significance level	مستوى المعنوية
Simultaneous-equations	معادلات آنيـــة
Skewness	التـــواء
Specification of Model	تحديد النموذج
Standard deviation	انحراف معيارى
Statistic	إحصائية
Statistics	علم الإحصاء
Stochastic	عشوائى (احتمالى)
Stratified sampling	معاينة طبقية
Structural coefficients	معاملات هيكلية
Structural (behavioral) equations	معادلات هيكلية (سلوكية)
Symmetry	تماثل
Systematic sampling	معاينة منتظمة

(T)

Time-series analysis	تحليل السلاسل الزمنية
Type I Error	خطأ من النوع الأول
Type II Error	خطأ من النوع الثانى

(U)

Unbiased estimate	تقدير غير متحيز
Underidentified equation	معادلة ناقصة التمييز
Unexplained residual	بواقى غير مفسرة
Ungrouped data	بيانات غير مبوبة

(V)

Variable	متغـــير
Variance	تبـــاين
Variation, coefficient of	معامل الاختلاف
Weighted mean	متوسط مرجح

الفهرس الأبجــدى

ملاحظات

ملاحظات..